

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 04:43:04 ; Search time 93 Seconds
(without alignments)
9928.116 Million cell updates/sec

Title: US-10-069-062-6
Perfect score: 1299
Sequence: 1 atgcacaaagcatttagtcgc.....aagactatagtttataa 1299

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUTS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1286.2	99.0	1299	4	US-09-792-024-42 Sequence 42, Appl
2	1161	89.4	1230	4	US-09-248-796A-3819 Sequence 3819, Ap
3	678.6	52.2	765	4	US-09-248-796A-3820 Sequence 3820, Ap
4	132	10.2	1356	4	US-09-614-221A-513 Sequence 513, App
5	65	5.0	90	4	US-09-792-024-348 Sequence 348, App
6	53.8	4.1	7218	1	US-08-232-463-14 Sequence 14, Appl
7	46.2	3.6	1141	4	US-09-806-708B-22 Sequence 22, Appl
8	45.6	3.5	1141	4	US-09-806-708B-22 Sequence 22, Appl
9	44.6	3.4	1356	4	US-09-540-236-117 Sequence 117, Ap
10	44.2	3.4	832	4	US-09-621-976-2813 Sequence 2813, Ap
11	43	3.3	119211	4	US-09-596-002-40 Sequence 40, Appl
12	42.8	3.3	2296	4	US-09-888-863A-1 Sequence 1, Appl
13	42.4	3.3	832	4	US-09-621-976-2813 Sequence 2813, Ap
14	40.4	3.1	728	4	US-09-988-863A-4 Sequence 4, Appl
15	38.8	3.0	4106	1	US-08-434-823-1 Sequence 1, Appl
16	38.8	3.0	14078	4	US-08-457-366-1 Sequence 191, App
17	38.6	3.0	14078	4	US-08-457-366-1 Sequence 191, App
18	38.6	3.0	14078	4	US-08-457-366-1 Sequence 191, App
19	38.4	3.0	1455	4	US-09-248-796A-2077 Sequence 2077, Ap
20	37.4	2.9	2214	4	US-09-248-796A-216 Sequence 216, App
21	37.4	2.9	762	4	US-09-248-796A-13337 Sequence 13337, A
22	37.4	2.9	1550	4	US-09-601-198-105 Sequence 105, App
23	37.4	2.9	3037	4	US-09-710-279-3904 Sequence 3904, Ap
24	37.4	2.9	4020	4	US-09-710-279-3904 Sequence 3904, Ap
25	37.2	2.9	2815	1	US-07-671-817A-2 Sequence 2, Appl
26	37.2	2.9	3444	1	US-08-349-867-22 Sequence 22, Appl
27	37.2	2.9	3444	1	US-08-349-867-22 Sequence 26, Appl

28	37.2	2.9	3444	1	US-08-239-476-22 Sequence 22, Appl
29	37.2	2.9	3444	1	US-08-239-476-26 Sequence 26, Appl
30	37.2	2.9	3444	1	US-08-598-305A-22 Sequence 22, Appl
31	37.2	2.9	3444	1	US-08-598-305A-26 Sequence 26, Appl
32	37.2	2.9	3444	2	US-08-639-923A-22 Sequence 22, Appl
33	37.2	2.9	3444	2	US-08-639-923A-26 Sequence 26, Appl
34	37.2	2.9	3444	5	PCT-US95-05431-22 Sequence 22, Appl
35	37.2	2.9	3450	1	US-08-349-867-20 Sequence 20, Appl
36	37.2	2.9	3450	1	US-08-349-867-20 Sequence 20, Appl
37	37.2	2.9	3450	1	US-08-239-476-20 Sequence 20, Appl
38	37.2	2.9	3450	1	US-08-598-305A-20 Sequence 20, Appl
39	37.2	2.9	3450	2	US-08-639-923A-20 Sequence 20, Appl
40	37.2	2.9	3450	2	PCT-US95-05431-20 Sequence 20, Appl
41	37.2	2.9	3465	1	US-08-349-867-18 Sequence 18, Appl
42	37.2	2.9	3465	1	US-08-239-476-18 Sequence 18, Appl
43	37.2	2.9	3465	1	US-08-598-305A-18 Sequence 18, Appl
44	37.2	2.9	3465	2	US-08-639-923A-18 Sequence 18, Appl
45	37.2	2.9	3465	5	PCT-US95-05431-18 Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-792-024-42
; Sequence 42, Application US/09792024
; Patent No. 6783985
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug
; FILE REFERENCE: 10182-004-999
; CURRENT APPLICATION NUMBER: US/09/792,024
; NUMBER OF SEQ ID NOS: 490
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-792-024-42

Query Match 99.0%; Score 1286.2; DB 4; Length 1299;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 ATGTCAAAAGCATTTAGTGCACCTGGAAAGATTTCTGTCGATGATTTGGTTCTT 60
1 ATGTCAAAAGCATTTAGTGCACCTGGAAAGATTTCTGTCGATGATTTGGTTCTT 60
61 GAGCCATTTATGATGCTTATGTCAGACATTTGTCATCAAGATGATGATGATTTA 120
61 GAGCCATTTATGATGCTTATGTCAGACATTTGTCATCAAGATGATGATGATTTA 120
121 CCAAAAGGACCAAGTTGAAAGATTTGATCAAAATTTCTTACCCCAATTTGCAAC 180
121 CCAAAAGGACCAAGTTGAAAGATTTGATCAAAATTTCTTACCCCAATTTGCAAC 180
181 GGAGATGGGATATATCATATCAATCAATACAGAGAGCCAGAGATTCAGTCAAGC 240
181 GGAGATGGGATATATCATATCAATCAATACAGAGAGCCAGAGATTCAGTCAAGC 240
241 ATTAATTCATTTTATGAGCAACTATATTCATGTTTATGATTAATTCAGCAGCA 300
241 ATTAATTCATTTTATGAGCAACTATATTCATGTTTATGATTAATTCAGCAGCA 300
301 GATTTGATTTGAAATCAATCAATTTACTGAGACCTGATATTCATTCACAAGAGTACT 360
301 GATTTGATTTGAAATCAATCAATTTACTGAGACCTGATATTCATTCACAAGAGTACT 360

QY 361 GAAACCAAGACATCTCTGAAATGAGAAAAACATTTCTTTACCATCTCTGTCATTACC 420
DB 361 GAAACCAAGACATCTCTGAAATGAGAAAAACATTTCTTTACCATCTCTGTCATTACC 420
QY 421 GAAGTGGAAAAAGACCGGATTAAGTTCAATGGCAGAGATTAAGTTCAGTTTGGCACAAGT 480
DB 421 GAAGTGGAAAAAGACCGGATTAAGTTCAATGGCAGAGATTAAGTTCAGTTTGGCACAAGT 480
QY 481 TTATTTATCCCATTTTATCCCAATGTTATCAGTACGAAATTAAGTATTTTGCACAAGT 540
DB 481 TTATTTATCCCATTTTATCCCAATGTTATCAGTACGAAATTAAGTATTTTGCACAAGT 540
QY 541 GCACACATTTGACATTTGTTATGCCCAAAAAAGATAGATCTGGTTTGAATTTGCAACT 600
DB 541 GCACACATTTGACATTTGTTATGCCCAAAAAAGATAGATCTGGTTTGAATTTGCAACT 600
QY 601 GCAATTTATGCTCTGATGTTATGATAGAAAGATTTCAAGCAGCTTTGATTAATGACGTTT 660
DB 601 GCAATTTATGCTCTGATGTTATGATAGAAAGATTTCAAGCAGCTTTGATTAATGACGTTT 660
QY 661 CAGGTTCTAGAAAGTATCTCTGAGAAAGTTCCCAAGATTTGAAAAATTTGATTTGAAGT 720
DB 661 CAGGTTCTAGAAAGTATCTCTGAGAAAGTTCCCAAGATTTGAAAAATTTGATTTGAAGT 720
QY 721 AACTGGGAATTCGAATGAAATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 780
DB 721 AACTGGGAATTCGAATGAAATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 780
QY 781 GACGTCAAGAGTGGCTCAGAAACACCAATTTGATCAAGTATTCATTCATTCATTCATTCAT 840
DB 781 GACGTCAAGAGTGGCTCAGAAACACCAATTTGATCAAGTATTCATTCATTCATTCATTCAT 840
QY 841 GAAAAACCCAGAAAGAAAGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 900
DB 841 GAAAAACCCAGAAAGAAAGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 900
QY 901 ATGAAGAAATTTGAGGAAATGCTGTAATAATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 ATGAAGAAATTTGAGGAAATGCTGTAATAATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 GAGTTAGATCAATCTGTTGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 GAGTTAGATCAATCTGTTGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 CAAGCTTTAACAAGAAATCAAGAGTTTCAATTTGAATCTGATGATGATGATGATGATGAT 1080
DB 1021 CAAGCTTTAACAAGAAATCAAGAGTTTCAATTTGAATCTGATGATGATGATGATGATGAT 1080
QY 1081 GACCGTTCTCAAGAGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
DB 1081 GACCGTTCTCAAGAGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
QY 1141 GATGCAATAGCTGTATTAAGTTGAAATCAAGTGGAAATTTTAAAGAGAAAACTCTT 1200
DB 1141 GATGCAATAGCTGTATTAAGTTGAAATCAAGTGGAAATTTTAAAGAGAAAACTCTT 1200
QY 1201 GAAATTCAGATTTATTTTATTAATGTTTATCTGGTTGATTTGAAAGCAACAGAGT 1260
DB 1201 GAAATTCAGATTTATTTTATTAATGTTTATCTGGTTGATTTGAAAGCAACAGAGT 1260
QY 1261 GTACTTGAAGAAAAACAGAAAGATATATAGTTTATTA 1299
DB 1261 GTACTTGAAGAAAAACAGAAAGATATATAGTTTATTA 1299

RESULT 2

US-09-248-796A-3819
; Sequence 3819, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 3819
LENGTH: 1230
TYPE: DNA
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: unsure
LOCATION: (1194)
OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
US-09-248-796A-3819
Query Match 89.4%; Score 1161; DB 4; Length 1230;
Best Local Similarity 98.2%; Pred. No. 3.3e-305;
Matches 1173; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 7 AAAGCATTTAGTGCACCTCGAAAAAGCATTTCTGCTGGTGAATTTGGTCTTGAGCCA 66
DB 34 AAAGCATTTAGTGCACCTCGAAAAAGCATTTCTGCTGGTGAATTTGGTCTTGAGCCA 93
QY 67 ATTTATGATGCTTATGATGACAGCATTTGTCATCAGAAATGATGATGATGATGATGATGATGAT 126
DB 94 ATTTATGATGCTTATGATGACAGCATTTGTCATCAGAAATGATGATGATGATGATGATGATGAT 153
QY 127 GGAACAGTTTGAAGAAATCTAGAAATCAAAATTTCTTACCCCAATTTGCAACGAGAA 186
DB 154 GGAACAGTTTGAAGAAATCTAGAAATCAAAATTTCTTACCCCAATTTGCAACGAGAA 213
QY 187 TGGAAATTCATATATCATATCAATATCAGAAAGCCGAGAAAGTTCAAGTCAAGTAAAT 246
DB 214 TGGAAATTCATATATCATATCAATATCAGAAAGCCGAGAAAGTTCAAGTCAAGTAAAT 273
QY 247 CCATTTTATGAGGCAATATATATCATGTTTATGATTTATTAATTAATTAATTAATTAATTAAT 306
DB 274 CCATTTTATGAGGCAATATATATCATGTTTATGATTTATTAATTAATTAATTAATTAATTAAT 333
QY 307 GATCTTGAATCATATTTATCTCAGACCTCGATATCATATTAATTAATTAATTAATTAATTAAT 366
DB 334 GATCTTGAATCATATTTATCTCAGACCTCGATATCATATTAATTAATTAATTAATTAATTAAT 393
QY 367 AAAGCATCTCGAATGAGAAAGAAACATTTCTTAAACATCTGTCATTAACGAGATG 426
DB 394 AAAGCATCTCTTAATGAGAAAGAAACATTTCTTAAACATCTGTCATTAACGAGATG 453
QY 427 GAAAAACCCGATTAAGTTCAATCGGAGATTAAGTTCAGTTGTTGCAACAAGTTATTA 486
DB 454 GAAAAACCCGATTAAGTTCAATCGGAGATTAAGTTCAGTTGTTGCAACAAGTTATTA 513
QY 487 TCCCATTTTATCCCAATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
DB 514 TCCCATTTTATCCCAATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573
QY 547 ATTCACATTTGTTATGTCGCAAAAAAGATGAGATCTGGTTTGAATTTGCAACGTTGCA 606
DB 574 ATTCACATTTGTTATGTCGCAAAAAAGATGAGATCTGGTTTGAATTTGCAACGTTGCA 633
QY 607 TATGCTGATTTGATATTAAGATTTCAAGCAGCTTTGATTAATGATGATGATGATGATGATGAT 666
DB 634 TATGCTGATTTGATATTAAGATTTCAAGCAGCTTTGATTAATGATGATGATGATGATGATGAT 693
QY 667 CTAAAAAGTATCTTGAAGATTTCCCAAGAGTTGAAAAATTTGATTAAGTAACTGG 726
DB 694 CTAAAAAGTATCTTGAAGATTTCCCAAGAGTTGAAAAATTTGATTAAGTAACTGG 753
QY 727 GAATTCAAACATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
DB 754 GAATTCAAACATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 813

649 AATGACGTGTTTCAGGTTCTAGAAAGTGATCCTCTACACCTCAGGATTTCTG


```
Db 1481 AATTACCTATGCAAGTACTTAAAGATGAGAAATTTGGTACRRRRRRRRRR 1422
Qy 709 TTGATGAAAGTACTGGAAATTCAAACATGAAAGATGATACCATGCAATCAAG 768
Db 1421 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1362
Qy 769 TTATTAATGAGTCAAGGCTGCAAGGCTGAGAAACACCAAAATGGTATCAGAGTAC 828
Db 1361 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1302
Qy 829 CAATGAAAAAGAAAAGCCAGAGAAAGCTCTGTTGTATGACCACTTAATAGTCC 888
Db 1301 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1242
Qy 889 AATTTACAGTTATGAGAAATGAGGAAATGCGGAAATATGACTCAGACCCAGAG 948
Db 1241 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1182
Qy 949 ACTTATATTTAAAGATTGATCATCTCTGTGAGCCCTTGAAGTATGCAATGATC 1008
Db 1181 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1122
Qy 1009 AGAAAAGGTTTCAAGCATTAACAAAAATCAGAG 1045
Db 1121 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1085
```

RESULT 7

```
US-09-806-708B-22/C
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22
```

Query Match 3.6%; Score 46.2; DB 4; Length 1141;
Best Local Similarity 13.5%; Pred. No. 0.0087;
Matches 94; Conservative 220; Mismatches 374; Indels 7; Gaps 1;

```
Qy 581 CTGGTTGATGTCGCAACTGCAATTATGCTGATATATAGAGATTCAGCCAG 640
Db 1136 YTCGTTTCTTCTTTCATNNNNNNNNNGMCDNNRMDATKSAIGTAMWTHAKRGATMCT 1077
Qy 641 CTTGATAAATGACGTTTCAAGTTCTAGAAAGTATGATCCCAAGAGT 700
Db 1076 WYWTGTNRRCWRTYAMRTMYRNSANMWSCATYBMMWTKWYATKYRTAMVMMCMWRNN 1017
Qy 701 TGAATAAATGATGTAAGTAACTGGAATTCAAACATGAAAGATGATACCTACG 760
Db 1016 NNMCAATNGAKSCATNNNAWYATTRWAAAYAAAKMAMAGNNMRMGAAAGNMGCAAAA 957
Qy 761 GAATCAAGTTATTAATGAGTGAAGTCAAGGCTGAGAAAACCCCAATTTGTTAC 820
Db 956 TMGBWADTDAGKCMNNNNNNNTTVDYRMMAMKAKNNNNNNAYTACTNRAATNNKATTHM 897
Qy 821 GAGTACTCCATGAGAAAAGAAAAGCCGAAAGAAAGCTCTGTGTGTATGACCAAGCTTA 880
```

```
Db 896 KWTGASKRRTTRHHTTCRTKTYNNNNNNNAETVYVYHHAARMNANWTRTNNNNNN 837
Qy 881 ATAGTGCATTTTACAGTTTATGAAGAATTGAGGAAATGCGTGAATAATGACCTCAG 940
Db 836 NNACATRTWABWGHSHCMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 777
Qy 941 ACCGAGACTTATATTAAGATGATGATCATCTGTTGAGCCTTGAAGCTTGGCATTA 1000
Db 776 TGNNTWMAAVTTHTHDWCYKTMWNTWYMDMTMB-----TTTBNMTSTNNNNNN 724
Qy 1001 AGAATCATGAAAAAGGTTTCAAGCATTAACCAAAAATCAGAGTCCATTTGAACCTG 1060
Db 723 NNMWACTNNNNNNMMKAYAAHTNNWGCWNNNTDARFTNNNTTMMRRMWTNKTWYST 664
Qy 1061 ATGTCAAACCCAGTTGTTGACCGCTGTGCAAGAGATTCCTGTTGTGTGAGTGAG 1120
Db 663 RHHHTTGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 604
Qy 1121 TTCAGAGTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1180
Db 603 VDVWADSWVWYANMRCADVTYTNNTYCKSYASVYWSNNAMWYRRYSARNWSMAR 544
Qy 1181 ATTTAAGCAAAAATCTTGAATAATCCAGATTAATTTATATATGTTTACTGGTGTGAT 1240
Db 543 WTRNNNNMMWSGEBRWRWAGTMMWRMNNNNNTDRTYMMWKRABRTTYVDSMCAKSW 484
Qy 1241 TGAAGAGCAACAGAGGTTACTTGAAGAAA 1275
Db 483 RNNWRNMMKMMMAANNNDGAMDHTYMGNNNTMM 449
```

RESULT 8

```
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22
```

Query Match 3.5%; Score 45.6; DB 4; Length 1141;
Best Local Similarity 9.5%; Pred. No. 0.013;
Matches 83; Conservative 325; Mismatches 460; Indels 2; Gaps 1;

```
Qy 252 TTTAGAGCAACTATATATCATGTTTATGCTTATATTAACAAGCAGACGATTTGATCT 311
Db 58 KWTWABMYCKYRBYNNNSRMWKGWYKKWYBNCANNTSBRVYARBMKDKTAYBMTMTNK 117
Qy 312 TGAATTCATCTTATCTCAGACCTCGATATCTTCAACAAGAGTACTGAAACCAAGAC 371
Db 118 WGTGTMRHYMRMBDTVDHRYVTANNNNAWTTTCMDKDCKRTYRMMKNNNATGMD 177
Qy 372 ATCTCGAATGAGAAAAAATCTTCTTATCATCTGTCGTCATTAACGAGTGAAGAAA 431
Db 178 DTKYHMMNNNGCBYTYMWRVYKTDSDSBKKNMYGMBWKKWSYVTTYWYVWDDMCGRK 237
Qy 432 GACCGAATTAAGTTTCATGCGCAGATTAAGTGTCAAGTTGTGCAAGTTATTCCTCA 491
```

Query	Best Local Similarity	Score	DB	Length
Matches 131; Conservative 0; Mismatches 144; Indels 0; Gaps 0;	3.4%;	44.6;	DB 4;	1356;
	47.6%;	Pred. No. 0.026;		
QY	1019	TACAAAGATTAAACCAAAATCAGAGTTCAAATTGAACTGTATGCCAAACCAATTTT	1078	
DB	545	TCCAAGTTCAAGCAGCCAAACCGAAGCTCAAATATGCTTTGGTATGATTCAGGTCTTACA	604	

```

RESULT 10
US-09-621-976-2813
: Sequence 2813, Application US/09621976
: Patent NO. 6639063
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: ESTs and Encoded Human Proteins
: FILE REFERENCE: GENSET_054PR2
: CURRENT APPLICATION NUMBER: US/09/621,976
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 19335
: SOFTWARE: Patent.pm
: SEQ ID NO 2813
: LENGTH: 832
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 235..399
: US-09-621-976-2813

```

	Query Match	3.4%;	Score 44.2;	DB 4;	Length 832;
	Best Local Similarity	15.28;	Frid. No. 0.026;		
	Matches	60;	Conservative 164;	Mismatches 167;	Indels 4; Gaps 1
339	ATATCAATTCACAAAGAGTACTGAAACCAAGACATCGTGATGTAGAGAAAAACATTCT	398			
11	WYAKCWTWKMSWSYMMWYMKTYMKTYMRKKKKAMKWTWTWYWRVAMNGTYKK	70			
399	TTACCAATTCGCGCATTAACCGAAGTGAAGAAACCGGATTTGATTCATCGCAGATT	458			
71	KAMCRKTKTKKKKKGGTMMWYMGWRSYVAWMTRTWTGYAYRSMMYWRIRCKKAYI	130			
459	AGTGCAGTTGTTGGCACAAAGTTTATATATCCCATTTTATCCCAATGTTACGTACGAA	518			
131	RKTYCYSKGMTWMBRWKKAATTTWTKCTYTAATRTWMMNC---WTKRPAASWYCMW	186			
519	TAAAGATATTTTGCACAAAGTTGCACAGATTGCACATTTGATGATCCCAAAAAAGATAG	578			
187	WKAARWSTWRSRSYASARSARCCYSCSWGANSWYMRMRMRMAAGAKMARAS	246			
579	ATCTGGGTTTATGTTGGCACTGCAATTTATATGCTCATTTGATATAGAAGATTACGCC	638			
247	CMRRKRYAGSKTGYKSSMMCMWTSMKYCYTKARWTDGYCYRKGGMWGRGRWYASKYV	306			
639	AGCTTTGATTAATACGTTTTCAGTTTCAGTTTCAGTAAGATCTCGAAGAAGTCCCCACAGA	698			
307	WKRMMWCMARMYRSTGTBRASMMWRMYVMMMKMKYMAAARMAKMMWMAWRPACAA	366			
699	GTTCAAAAATATGTTTAAAGTAACTCGGCAATTCA	733			
367	AATATAATTTATTTATAGCTACAAATCTTGTGAACTTTA	401			

;; CURRENT FILING DATE: 2001-11-21
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO: 4
;; LENGTH: 728
;; TYPE: DNA
;; ORGANISM: Gossypium hirsutum
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: 715
;; OTHER INFORMATION: n can be any nucleotide
US-09-988-863A-4

Query Match

Best Local Similarity 3.1%; Score 40.4; DB 4; Length 728;
Matches 62; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 514 ACGAATTAAGATATTTTGCACAAGTTGCACAGATTGCACATTGTTAGCCCCAAAAAG 573
DB 562 ACAGATCTGATTTGTGCAATGATGAGCTCAAGGCCACCTGATTTGCCCAAGTAA 621
QY 574 ATGAGATCTGGTTTATGTTGCAACTGCATTATAGG 611
DB 622 GTTGGCAGTGGCTTTGATGTCAGTTCTGCTGTATGG 659

RESULT 15

US-08-434-823-1
; Sequence 1, Application US/08434823
; Patent No. 5712248

GENERAL INFORMATION:

;; APPLICANT: Kalman, Sue S.
;; APPLICANT: Kiehne, Kristine L.
;; TITLE OF INVENTION: NOVEL INSECTICIDE PROTEIN AND GENE
;; NUMBER OF SEQUENCES: 13
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sandoz Agro, Inc.
;; STREET: 975 California Avenue
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/434,823
;; FILING DATE:
;; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/197,998
;; FILING DATE: 16-FEB-1994
;; APPLICATION NUMBER: US 08/102,316
;; FILING DATE: 05-AUG-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/004,474
;; FILING DATE: 14-JAN-1993
;; APPLICATION DATA:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Marcus-Wyner, Lynn
;; REGISTRATION NUMBER: 34,869
;; REFERENCE/DOCKET NUMBER: Z-702/CONT3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/857-1125
;; TELEFAX: 415/857-1125

INFORMATION FOR SEQ ID NO: 1:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4106 base pairs
;; TYPE: nucleic acid

;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 296..3826
;; OTHER INFORMATION: /codon_start= 296
US-08-434-823-1

Query Match

Best Local Similarity 3.0%; Score 38.8; DB 1; Length 4106;
Matches 58; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 903 GAAGATTTAGGGAATTCCTGAAATTAAGACTACAGCCAGAGACTTATATTAAAGA 962
DB 2929 GAAGAAATGAGAGACAAACGTGAAAAATTTGGAATGGGAAACAATATTTTATAAAGA 2988
QY 963 GTTAGATCTTCTGTGAGCCCTTGACTGT 992
DB 2989 GGCAAAAGATCTGTAGATGCTTTATTTGT 3018

RESULT 16

US-08-457-366-1
; Sequence 1, Application US/08457366
; Patent No. 5731194

GENERAL INFORMATION:

;; APPLICANT: Kalman, Sue S.
;; APPLICANT: Kiehne, Kristine L.
;; TITLE OF INVENTION: NOVEL INSECTICIDE PROTEIN AND GENE
;; NUMBER OF SEQUENCES: 13
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sandoz Agro, Inc.
;; STREET: 975 California Avenue
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/457,366
;; FILING DATE:
;; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/197,998
;; FILING DATE: 16-FEB-1994
;; APPLICATION NUMBER: US 08/102,316
;; FILING DATE: 05-AUG-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/004,474
;; FILING DATE: 14-JAN-1993
;; APPLICATION DATA:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Marcus-Wyner, Lynn
;; REGISTRATION NUMBER: 34,869
;; REFERENCE/DOCKET NUMBER: Z-702/CONT3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/857-1125
;; TELEFAX: 415/857-1125

INFORMATION FOR SEQ ID NO: 1:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4106 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 296..3826

OTHER INFORMATION: /codon_start= 296
US-08-457-366-1

Query Match 3.0%; Score 38.8; DB 1; Length 4106;
Best Local Similarity 64.4%; Pred. No. 1.6;
Matches 58; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 903 GAAGAAATGAGGAAATCGTGAATAATACGACTGAGACCGAGACTTAATTAAGA 962
DB 2929 GAAGAAATGAGGAGCAAAACGTGAATAATGAGATGGAACAAATATTGTTATAAGA 2988

QY 963 GTTAATCATCTGTTGACCTTTGACTGT 992
DB 2989 GCGAAAGATCTGTAGATGCTTAATTGT 3018

RESULT 17
US-08-956-171E-191/c

; Sequence 191, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:

APPLICANT: Charles Kunach

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 191:

SEQUENCE CHARACTERISTICS:

LENGTH: 14078 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 191:

QY 65 CAATTATGATGCTTATGTGACAGCATTTGTCATCAAGATGATGATTAACACCA 124

DB 2341 TGGCTTCAATCAAAATGTTTCATGCTGTATTCACGATACCTGTACAGCATCTCTAT 2282

QY 125 AAGAACCACTTT 137

DB 2281 ATTTAAACATATTT 2269

RESULT 18

US-08-781-986A-191/c

; Sequence 191, Application US/08781986A

; Patent No. 6737248

; GENERAL INFORMATION:

APPLICANT: Charles Kunach

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 191:

SEQUENCE CHARACTERISTICS:

LENGTH: 14078 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-191

Query Match 3.0%; Score 38.6; DB 4; Length 14078;

Best Local Similarity 55.6%; Pred. No. 3.2;

Matches 74; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 5 CAAGACATTTAGTCACCTGGAAGAAAGATTTCTGCTGCGAGATATTTGTTCTTGACG 64

DB 2401 CATTGCAATGCTGTGTATGATTAATGCAATTTGTTCTGATGCTTAATTAATCTTGTC 2342

QY 65 CAATTATGATGCTTATGTGACAGCATTTGTCATCAAGATGATGATTAACACCA 124

DB 2341 TGGCTTCAATCAAAATGTTTCATGCTGTATTCACGATACCTGTACAGCATCTCTAT 2282

QY 125 AAGAACCACTTT 137

DB 2281 ATTTAAACATATTT 2269

RESULT 19

US-09-248-796A-2077/c

; Sequence 2077, Application US/09248796A

; Patent No. 6747137

```

; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2077
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-2077
```

```

Query Match
Best Local Similarity 3.0%; Score 38.4; DB 4; Length 1455;
Matches 99; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
```

```

QY 70 TATGATCTTATGTGACAGCATTTGTCATCAGCAATGCATGCACTTAAACACCAAGGA 129
DB 533 TATTTTCTGATGTAAATATTAACACAGAAATAGCCGAGCAATCTAAGAAATGA 474
QY 130 ACCAGTTTGAAAGAAATCTAGAAATCAAAATTTCTTACCCCAATTTGCAACGAGAAATGG 189
DB 473 ACAATGCAATACCTTTTCTGCAATGATTTTAATACATTCATGACCATTAATCTC 414
QY 190 GAATATCAATATCATCAATAACAGAGAAAGCCAGAGAAAGTTCAGCGATTAATCCA 249
DB 413 AAATCATTAATATTTCTTCTTCAGATTAATCCAGCGAATTAATCAACCAATTAACA 354
QY 250 TTTTGAAGGCACTATATTT 269
DB 353 TTTCTGTCGCTCCATCATTT 334
```

RESULT 20

```

US-09-248-796A-216
; Sequence 216, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 216
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (179)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-248-796A-216
```

```

Query Match
Best Local Similarity 2.9%; Score 37.8; DB 4; Length 2214;
Matches 72; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```

```

QY 929 AATAGACTGACACCCAGAGCTTATATTAAGAGTTAGATCAATCGTTGAGCCCTTGA 988
DB 425 AATAGCAACAGAGAGAAAGTTATGCTGATTTGATTAATGACATATCAACCCCTTGA 484
```

```

QY 989 CTGTTGGATTAAGAAATCAGAAAAAGGTTTACAACTTAACACAAAATCAGAGTTC 1048
DB 485 CTGATGATTTATTAATCAATTAATAGTAGCTTAAGAAAGACTTAACAAAGATGAAGATTG 544
QY 1049 CAATTGAC 1057
DB 545 CAATGAC 553
```

RESULT 21

```

US-09-248-796A-13337
; Sequence 13337, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 13337
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-13337
```

```

Query Match
Best Local Similarity 2.9%; Score 37.4; DB 4; Length 762;
Matches 116; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
```

```

QY 684 GAAATCCCAAGAGTTGAAAAAATTGATTAAGTAATCGGAATTCAAACATGAAAG 743
DB 156 GAATACCTTTTACAGTCGAAACAAACCGCTAAGCAAAAGAGATTGAATTCAAAG 215
QY 744 ATGTACATTAACATACGAAATCAATTAATTAATGGTGCATCAAGGTGCTCAGAAC 803
DB 216 AGAGCAGTAGACACCGATTTGAATTCGCTGTCGCAATACCAAAATGTCATGACAC 275
QY 804 ACCCAATGTTATACAGAGTACTCAATGAAAAAGAAAAAGCCAGAGAAAGCTCTGT 863
DB 276 ATCAAAATTAATCAAAAAATCAACACAGATTTTGAATATATCACTGAATGACAT 335
QY 864 TGTGATGACAGCTTAATAGGCCAATTAACATTAATGAGGAATGAGGAATGCG 923
DB 336 GGTAGTATGTTTATAGCAATTTCTTACGTAACAACTACAAATGAATTCATCAATCCA 395
QY 924 TGAATAA 930
DB 396 TCTTAAA 402
```

RESULT 22

```

US-09-601-198-105/c
; Sequence 105, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Casseil, Gail H.
; APPLICANT: Chen, Elison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
```

;; PRIOR FILING DATE: 1998-01-30
;; NUMBER OF SEQ ID NOS: 181
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 105
;; LENGTH: 1550
;; TYPE: DNA
;; ORGANISM: Ureaplasma urealyticum
US-09-601-198-105

Query Match 2.9%; Score 37.4; DB 4; Length 1550;
Best Local Similarity 44.0%; Pred. No. 2.4; Indels 0; Gaps 0;
Matches 158; Conservative 0; Mismatches 201;

QY 12 ATTAGTCGACCTGGAAAAGCATTTCTGCTGGTGAATTTGCTTCTGACCAATTTA 71
DB 372 ATTTAAACACTGTTATTAATAAACAAGTTGATGTTATTTAGATGTTAAAAAAGAAC 313
QY 72 TGATGCTTATGTCAGCAGCTTGTATCAGCAATGTCAGTATTAACCAAAAGAAC 131
DB 312 AGATATATGATGATCTTTTATTAATTAATAAAGATTCAGAGATTAATGGCGAAC 253
QY 132 CAGTTTGAAGAATCTAGATCAAAATTTCTCACCCTAATTGCAAGGAAATGGGA 191
DB 252 TTTTATTAATCAGCTAAAGGTACATATTTATTTCTATCTGTTGCCAAATCTACATATAT 193
QY 192 ATATCATATCATCATCAATATACAGAGAGCCAGAGAGTTCACTACGCATTAATCATTT 251
DB 192 TTCTACACATCTTCAAGAAATTTAAAGAAATATGGATTAATCTAATTAATCAATATGA 133
QY 252 TTTAGAGGCACTATATATTCATGCTTTTACCTTATATTAATCAACGACCGAAGCATTTGATCT 311
DB 132 AAAACAGACATTAATTTGGTCTTTTATCTTATGAGAGCAAAAGTAAAGATGTGG 73
QY 312 TGAATCATCATTTTACTGAGCCCTGATATCATTTACAGAAAGATCTGAAACCAAGA 370
DB 72 ATATTCGTAATAATGCTACACGCTCTAAAGAAATTTCTATGCAAAACATGAATTAAGA 14

RESULT 23

US-09-710-279-3904/c
; Sequence 3904; Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 3904
; LENGTH: 3037
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-3904

Query Match 2.9%; Score 37.4; DB 4; Length 3037;
Best Local Similarity 57.1%; Pred. No. 3.3; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 51;

QY 733 AAACATGAAAGATGTACATTAACGATCAAGTTATTAATGAGTACGTCAGGCT 792
DB 264 AAAAATGAAAGATGTATTAATAATGTAATTAACATAGTATTAATCTCTATAGAGG 205
QY 793 GGCTCAGAAACACCCAAATTTGGTATACAGATCTCAATGAAAAAGAAAGCCAGA 851
DB 204 GACTTGTTAAATAAGTTATTAATGATAGTTCTCAATGGTCAAAATTAAGCCAGA 146

RESULT 24

US-09-710-279-3434/c
; Sequence 3434; Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 3434
; LENGTH: 4020
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-3434

Query Match 2.9%; Score 37.4; DB 4; Length 4020;
Best Local Similarity 57.1%; Pred. No. 3.8;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 733 AAACATGAAAGATGTACATTAACGATCAAGTTATTAATGAGTACGTCAGGCT 792
DB 3089 AAAAATGAAAGATGTATTAATAATTTGTAATTAACATAGTATTTCTCTATAGAGG 3030
QY 793 GGCTCAGAAACACCCAAATTTGGTATCAGAGTACTCCATGAAAAAGAAAGCCAGA 851
DB 3029 GACTTGTTAAATAAGTTATTAATGATAGTTCTCAATGGTCAAAATTAAGCCAGA 2971

RESULT 25

US-07-671-817A-2
; Sequence 2; Application US/07671817A
; Patent No. 5424409
; GENERAL INFORMATION:
; APPLICANT: Ely, Susan
; TITLE OF INVENTION: DNA constructs
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby and Cushman
; STREET: Eleventh floor, 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-3601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/671,817A
; FILING DATE: 19910401
; CLASSIFICATION: 435
; APPLICATION DATA:
; APPLICATION NUMBER: GB 8823068.5
; FILING DATE: 30-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB89/01157
; FILING DATE: 29-SEP-1989
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 2815 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-671-817A-2

Query Match
Best Local Similarity 2.9%; Score 37.2; DB 1; Length 2815;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGGAATTGAGGGAATGCGTGAATAATGAGCTGAGACCCGAGAGCTTATATTAAAGA 962
DB 2642 GAAAAAATGAGAGACAAACGTGAAAAATGGAATGGAAACAAATATTGTTATAAAGA 2701
QY 963 GTTAGATCATCTTCTGTAGCCCTTTGACTGT 992
DB 2702 GGCAGAAAGATCTGTAGATGCTTTATTGT 2731

RESULT 26
US-08-349-867-22
Sequence 22, Application US/08349867
Patent No. 5508264
GENERAL INFORMATION:
APPLICANT: Bradfisch, Gregory A.
APPLICANT: Thompson, Mark
TITLE OF INVENTION: No. 5508264e1 Pesticidal Compositions
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/349,867
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA86
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-349-867-22

Query Match
Best Local Similarity 2.9%; Score 37.2; DB 1; Length 3444;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGGAATTGAGGGAATGCGTGAATAATGAGCTGAGACCCGAGAGCTTATATTAAAGA 962
DB 2550 GAAAAAATGAGAGACAAACGTGAAAAATGGAATGGAAACAAATATTGTTATAAAGA 2609
QY 963 GTTAGATCATCTTCTGTAGCCCTTTGACTGT 992

DB 2610 GGCAGAAAGATCTGTAGATGCTTTATTGT 2639

RESULT 27
US-08-349-867-26
Sequence 26, Application US/08349867
Patent No. 5508264
GENERAL INFORMATION:
APPLICANT: Bradfisch, Gregory A.
APPLICANT: Thompson, Mark
TITLE OF INVENTION: No. 5508264e1 Pesticidal Compositions
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/349,867
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA86
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ. ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-349-867-26

Query Match
Best Local Similarity 2.9%; Score 37.2; DB 1; Length 3444;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGGAATTGAGGGAATGCGTGAATAATGAGCTGAGACCCGAGAGCTTATATTAAAGA 962
DB 2550 GAAAAAATGAGAGACAAACGTGAAAAATGGAATGGAAACAAATATTGTTATAAAGA 2609
QY 963 GTTAGATCATCTTCTGTAGCCCTTTGACTGT 992
DB 2610 GGCAGAAAGATCTGTAGATGCTTTATTGT 2639

RESULT 28
US-08-239-476-22
Sequence 22, Application US/08239476
Patent No. 5527883
GENERAL INFORMATION:
APPLICANT: Thompson, Mark
APPLICANT: Schwab, George E.
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida

Db 2550 GAAAAATGAGAGACAAACGTGAAAAATGGAATGGAACAAATTTGTTATAAGA 2609
QY 963 GTTAGATCATTCCTGTTGAGCCTTGACTGT 992
Db 2610 GGCAAAAGAAATCTGTAGATGCTTTATTGT 2639

RESULT 31

US-08-598-305A-26
Sequence 26, Application US/08598305A
Patent No. 5827514
GENERAL INFORMATION:
APPLICANT: BRADFISCH, Gregory A.
APPLICANT: THOMPSON, Mark
APPLICANT: SCHWAB, George E.
TITLE OF INVENTION: No. 5827514e1 Pesticidal Compositions
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/598.305A
FILING DATE: 08-FEB-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/349,867
FILING DATE: 06-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA86.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-598-305A-26

Query Match 2.9%; Score 37.2; DB 1; Length 3444;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGCAATTGAGGAAATCGTGAATAATGACACTCAGACCCAGACTTATATTAAAGA 962
Db 2550 GAAAAATGAGAGACAAACGTGAAAAATGGAATGGAACAAATTTGTTATAAGA 2609
QY 963 GTTAGATCATTCCTGTTGAGCCTTGACTGT 992
Db 2610 GGCAAAAGAAATCTGTAGATGCTTTATTGT 2639

RESULT 32

US-08-639-923A-22
Sequence 22, Application US/08639923A
Patent No. 5840554
GENERAL INFORMATION:
APPLICANT: Thompson, Mark

APPLICANT: Schwab, George E.
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
Pseudomonas fluorescens
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639.923A
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,476
FILING DATE: 06-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-639-923A-22

Query Match 2.9%; Score 37.2; DB 2; Length 3444;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGCAATTGAGGAAATCGTGAATAATGACACTCAGACCCAGACTTATATTAAAGA 962
Db 2550 GAAAAATGAGAGACAAACGTGAAAAATGGAATGGAACAAATTTGTTATAAGA 2609
QY 963 GTTAGATCATTCCTGTTGAGCCTTGACTGT 992
Db 2610 GGCAAAAGAAATCTGTAGATGCTTTATTGT 2639

RESULT 33

US-08-639-923A-26
Sequence 26, Application US/08639923A
Patent No. 5840554

GENERAL INFORMATION:
APPLICANT: Thompson, Mark
APPLICANT: Schwab, George E.
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
Pseudomonas fluorescens
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,923A
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,476
FILING DATE: 06-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-639-923A-26

Query Match 2.9%; Score 37.2; DB 2; Length 3444;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGAAATTGAGGAAATCGTGAATAATACGACCTCAGACCCAGACCTTATTATTAAGA 962
DB 2550 GAAAAAATGAGAGACCAACGTAATAATTTGGAATGCGAAACAAATATTGTTATTAAGA 2609
QY 963 GTTAGATCATCTCTGTTGAGCCTTTGACTGT 992
DB 2610 GGCAAAAGATCTGTAGATGCTTTATTGT 2639

RESULT 34

PCT-US95-05431-22
Sequence 22, Application PC/TUS9505431
GENERAL INFORMATION:
APPLICANT:
APPLICANT: Street address: 5501 Oberlin Drive
APPLICANT: City: San Diego
APPLICANT: State/Province: California
APPLICANT: Country: US
APPLICANT: Postal code/Zip: 92121
APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
APPLICANT: Telex number:
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
TITLE OF INVENTION: Pseudomonas fluorescens
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05431
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
PCT-US95-05431-26

REFERENCE/DOCKET NUMBER: MA83
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US95-05431-22

Query Match 2.9%; Score 37.2; DB 5; Length 3444;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGAAATTGAGGAAATCGTGAATAATACGACCTCAGACCCAGACCTTATTATTAAGA 962
DB 2550 GAAAAAATGAGAGACCAACGTAATAATTTGGAATGCGAAACAAATATTGTTATTAAGA 2609
QY 963 GTTAGATCATCTCTGTTGAGCCTTTGACTGT 992
DB 2610 GGCAAAAGATCTGTAGATGCTTTATTGT 2639

RESULT 35

PCT-US95-05431-26
Sequence 26, Application PC/TUS9505431
GENERAL INFORMATION:
APPLICANT:
APPLICANT: Street address: 5501 Oberlin Drive
APPLICANT: City: San Diego
APPLICANT: State/Province: California
APPLICANT: Country: US
APPLICANT: Postal code/Zip: 92121
APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
APPLICANT: Telex number:
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
TITLE OF INVENTION: Pseudomonas fluorescens
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05431
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US95-05431-26

Query Match 2.9%; Score 37.2; DB 5; Length 3444;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGAAATTGAGGAATGCGTGAATAATCGACTCAGACCCGAGAGCTTATTTAAAGA 962
DB 2550 GAAAAAATGAGAGACAAACGTGAAAAATTGGAATGGAAACAAATATTGTTATTAAGA 2609

QY 963 GTTAGATCATTCGTTGAGCCTTTGACTGT 992
DB 2610 GGCAAAAGATCTGTAGATGCTTTATTGT 2639

RESULT 36
US-08-349-867-20
; Sequence 20, Application US/08349867
; Patent No. 5508264
; GENERAL INFORMATION:
; APPLICANT: Bradfisch, Gregory A.
; APPLICANT: Thompson, Mark
; TITLE OF INVENTION: No. 5508264el Pesticidal Compositions
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,867
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA86
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-349-867-20

Query Match 2.9%; Score 37.2; DB 1; Length 3450;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGAAATTGAGGAATGCGTGAATAATCGACTCAGACCCGAGAGCTTATTTAAAGA 962
DB 2556 GAAAAAATGAGAGACAAACGTGAAAAATTGGAATGGAAACAAATATTGTTATTAAGA 2615

QY 963 GTTAGATCATTCGTTGAGCCTTTGACTGT 992
DB 2616 GGCAAAAGATCTGTAGATGCTTTATTGT 2645

RESULT 37
US-08-239-476-20
; Sequence 20, Application US/08239476
; Patent No. 5527883
; GENERAL INFORMATION:

APPLICANT: Thompson, Mark
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
; Pseudomonas fluorescens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/239,476
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA83
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-239-476-20

Query Match 2.9%; Score 37.2; DB 1; Length 3450;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGAAATTGAGGAATGCGTGAATAATCGACTCAGACCCGAGAGCTTATTTAAAGA 962
DB 2556 GAAAAAATGAGAGACAAACGTGAAAAATTGGAATGGAAACAAATATTGTTATTAAGA 2615

QY 963 GTTAGATCATTCGTTGAGCCTTTGACTGT 992
DB 2616 GGCAAAAGATCTGTAGATGCTTTATTGT 2645

RESULT 38
US-08-598-305A-20
; Sequence 20, Application US/08598305A
; Patent No. 5827514
; GENERAL INFORMATION:
; APPLICANT: BRADFISCH, Gregory A.
; APPLICANT: THOMPSON, Mark
; APPLICANT: SCHWAB, George E.
; TITLE OF INVENTION: No. 5827514el Pesticidal Compositions
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/598,305A
FILING DATE: 08-FEB-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/349,867
FILING DATE: 06-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA86.D1
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 3450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-598-305A-20

Query Match 2.9%; Score 37.2; DB 1; Length 3450;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGGAATTGAGGGAATCGTGAATAATACGACTCAGCCAGACTTATTTAAAGA 962
DB 2556 GAAAAAATGAGAGACAAACGTGAATAATTTGAATGGAAACAAATATTGTTATTAAGA 2615
QY 963 GTTAGATCATTTCTGTTGAGCCTTTGACTGT 992
DB 2616 GGCAAAAGATCTGTAGATGCTTTATTGT 2645

RESULT 39
US-08-639-923A-20
Sequence 20, Application US/08639923A
Patent No. 5840554
GENERAL INFORMATION:
APPLICANT: Thompson, Mark
APPLICANT: Schwab, George E.
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
TITLE OF INVENTION: Pseudomonas fluorescens
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,923A
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,476
FILING DATE: 06-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 3450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-639-923A-20

Query Match 2.9%; Score 37.2; DB 2; Length 3450;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 903 GAAGGAATTGAGGGAATCGTGAATAATACGACTCAGCCAGACTTATTTAAAGA 962
DB 2556 GAAAAAATGAGAGACAAACGTGAATAATTTGAATGGAAACAAATATTGTTATTAAGA 2615
QY 963 GTTAGATCATTTCTGTTGAGCCTTTGACTGT 992
DB 2616 GGCAAAAGATCTGTAGATGCTTTATTGT 2645

RESULT 40
PCT-US95-05431-20
Sequence 20, Application PC/TUS9505431
GENERAL INFORMATION:

APPLICANT:
APPLICANT: Street address: 5501 Oberlin Drive
APPLICANT: City: San Diego
APPLICANT: State/Province: California
APPLICANT: Country: US
APPLICANT: Postal code/Zip: 92121
APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
APPLICANT: Telex number:
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
TITLE OF INVENTION: Pseudomonas fluorescens
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05431
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA83
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 3450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US95-05431-20

Query Match 2.9%; Score 37.2; DB 5; Length 3450;
Best Local Similarity 63.3%; Pred. No. 4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 903 GAGGAAATTGAGGGAATGCGTGAATAATGACTCAGACCCAGAGACTTATATTAAAGA 962
 DB 2556 GAAAAAATGAGAGACAAACGTGAAAAATGSAATGGAAACAAATATTGTTATATAAGA 2615
 OY 963 GTTAGATCATTCCTGTGTGAGCCTTTGACTGT 992
 DB 2616 GGCAAAAGATCTGTAGATGCTTTATTGT 2645

Search completed: January 26, 2005, 06:40:22
 Job time : 96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 26, 2005, 03:39:29 ; Search time 3870 Seconds
(without alignments)
15873.209 Million cell updates/sec

Title: US-10-069-062-6

Perfect score: 1299

Sequence: 1 atgcacaaagatctatgctc.....aagactataggtttataa 1299

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1299	100.0	1299	6	AX087879 Sequence
2	1299	100.0	1299	6	AX087878 Sequence
3	1286.2	99.0	1299	6	AX240804 Sequence
4	1286.2	99.0	1299	6	AX488723 Sequence
5	521	40.1	577	6	AX087876 Sequence
6	470.4	35.2	547	6	AX087874 Sequence
7	323.6	24.9	11000	8	CR380239_14
8	201	15.5	11000	8	CR380952_03
9	158.8	12.2	314422	8	AE016898
10	148.4	11.4	110000	8	CR382123_06
11	132.2	10.2	2299	8	YSCERG8
12	132	10.2	1356	6	BD170980
13	132	10.2	1356	6	BD171098
14	132	10.2	40397	8	SC9959
15	101.6	7.8	110000	8	CR382131_07
16	86.6	6.7	42947	8	SPAC343
17	65	5.0	65	6	AX485723
18	65	5.0	90	6	AX241110
19	56	4.3	963	11	CNS06RLC

20	55.6	4.3	1037	11	CNS06C37	AL401081 T7 end of
21	53.8	4.1	7218	6	166494	166494 Sequence 14
22	51.8	4.0	253050	1	AP000984	AP000984 Sulfolobu
23	50.6	3.9	166918	9	AC009757	AC009757 Homo sapi
24	48.2	3.7	2000	6	AX655393	AX655393 Sequence
25	47.8	3.7	732	1	AY147182	AY147182 Streptoco
26	47.8	3.7	923	1	AY139423	AY139423 Streptoco
27	47.8	3.7	1288	1	SPEM41	X58178 S. pyogenes
28	46.4	3.6	949	1	AY139408	AY139408 Streptoco
29	46.4	3.6	951	1	AY139421	AY139421 Streptoco
30	46.2	3.6	1141	6	AX083744	AX083744 Sequence
31	46	3.5	220761	2	AC101700	AC101700 Mus muscu
32	45.8	3.5	233719	2	AC131407	AC131407 Rattus no
33	45.8	3.5	243143	2	AC125664	AC125664 Rattus no
34	45.6	3.5	1141	6	AX083744	AX083744 Sequence
35	45.4	3.5	581	8	AY279126	AY279126 Arisaema
36	45.4	3.5	105989	2	AC008121	AC008121 Homo sapi
37	45.4	3.5	110000	2	AC009727_2	Continuation (3 of
38	45	3.5	2000	6	AX655393	AX655393 Sequence
39	44.6	3.4	1356	6	AR450463	AR450463 Sequence
40	44.6	3.4	143298	8	AC134240	AC134240 Oryza sat
41	44.4	3.4	726	1	AY138855	AY138855 Streptoco
42	44.2	3.4	832	6	AR415176	AR415176 Sequence
43	44.2	3.4	832	6	AX972010	AX972010 Sequence
44	44.2	3.4	832	6	BD110729	BD110729 EST and e
45	43.8	3.4	96492	5	BX323547	BX323547 Zebrafish

ALIGNMENTS

RESULT 1	AX087879	1299 bp	DNA	linear	PAT 17-MAR-2001
LOCUS	AX087879	Sequence 6 from Patent WO0114533.			
DEFINITION	AX087879	Sequence 6 from Patent WO0114533.			
ACCESSION	AX087879.1	GI:13396872			
VERSION	AX087879.1	GI:13396872			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	Rosamond, J.D. and Schnell, N.F.				
TITLE	Phosphomethyltransferase kinase (pmk) gene (erg8) from candida albicans				
JOURNAL	Patent: WO 0114533-A 6 01-MAR-2001; Astrazeneca AB (SE)				
FEATURES					
source	1. 1299				
ORIGIN					
Query Match	100.0%; Score 1299; DB 6; Length 1299;				
Best Local Similarity	100.0%; Pred. No. 1.5e-280;				
Matches 1299; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 ATGTCAAAAGCATTTAGTGCACCTGGAAGCATTTCTTGCTGTGGATATTGTTCTT	60			
DB	1 ATGTCAAAAGCATTTAGTGCACCTGGAAGCATTTCTTGCTGTGGATATTGTTCTT	60			
QY	61 GAGCAATTATGATGCTTATGTGACAGATGTCATCGAATGATGATGATTTTAA	120			
DB	61 GAGCAATTATGATGCTTATGTGACAGATGTCATCGAATGATGATGATTTTAA	120			
QY	121 CCAAAAGGACGATTTGAAAGATCTGAATCAAAATTTCTTCAACCCCAATTTGCAAC	180			
DB	121 CCAAAAGGACGATTTGAAAGATCTGAATCAAAATTTCTTCAACCCCAATTTGCAAC	180			
QY	181 GAGGATGGGATTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	240			
DB	181 GAGGATGGGATTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	240			

QY 241 ATAAATCCATTTTATAGAGCACTATATTCATGTTTATAGCTTATATTTCAACCGACCGAA 300
 DB 241 ATAAATCCATTTTATAGAGCACTATATTCATGTTTATAGCTTATATTTCAACCGACCGAA 300
 QY 301 GCATTTATCTTGAATATCATCATTTACTCAGACCTTGGAATATCATTTCAAGAAAGATCT 360
 DB 301 GCATTTATCTTGAATATCATCATTTACTCAGACCTTGGAATATCATTTCAAGAAAGATCT 360
 QY 361 GAAACCAAGACATCTCGAATGAGGAAACAACTTCTTACCATTCCTGCTGCATTTACC 420
 DB 361 GAAACCAAGACATCTCGAATGAGGAAACAACTTCTTACCATTCCTGCTGCATTTACC 420
 QY 421 GAAATGAAAAAGACCGGATTAGGTTCAATCGGACAGATTAAGTCAAGTTGTTGCCACAAGT 480
 DB 421 GAAATGAAAAAGACCGGATTAGGTTCAATCGGACAGATTAAGTCAAGTTGTTGCCACAAGT 480
 QY 481 TTATTTATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGGCAACGTT 540
 DB 481 TTATTTATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGGCAACGTT 540
 QY 541 GCACAGATTGACATTTGTTATGCCCAAAAAAGATAGGATCTGGGTTGATGTTGCACT 600
 DB 541 GCACAGATTGACATTTGTTATGCCCAAAAAAGATAGGATCTGGGTTGATGTTGCACT 600
 QY 601 GCATTTATGCTGATGTTGATATAGAAAGATTTCAAGCTTGTGATTAATGACGTGTT 660
 DB 601 GCATTTATGCTGATGTTGATATAGAAAGATTTCAAGCTTGTGATTAATGACGTGTT 660
 QY 661 CAGGTTCTAGAAAGATGATCCTGAGAGTTCCCAAGAGTTGAAAAAATTTGATGAAAGT 720
 DB 661 CAGGTTCTAGAAAGATGATCCTGAGAGTTCCCAAGAGTTGAAAAAATTTGATGAAAGT 720
 QY 721 AACTGGAAATTCAAACATGAAAGATGATCATTACCATCAGATCAAGTTATTAATGAGT 780
 DB 721 AACTGGAAATTCAAACATGAAAGATGATCATTACCATCAGATCAAGTTATTAATGAGT 780
 QY 781 GACGTCAAGGTGGCTCAGAAACACCCAAATTTGATTCAGAGTACTCAATGAAAAAG 840
 DB 781 GACGTCAAGGTGGCTCAGAAACACCCAAATTTGATTCAGAGTACTCAATGAAAAAG 840
 QY 841 GAAAAACCAAGAAAGATCTGTTGATGATGACCAAGCTTAATAGGCCAATTTAGATTT 900
 DB 841 GAAAAACCAAGAAAGATCTGTTGATGATGACCAAGCTTAATAGGCCAATTTAGATTT 900
 QY 901 ATGAAAGAAATTCAGGAAATGCGTGAATAATACGACTCAGACCCAGAGCTTATATTA 960
 DB 901 ATGAAAGAAATTCAGGAAATGCGTGAATAATACGACTCAGACCCAGAGCTTATATTA 960
 QY 961 GAGTTAGATCATTTCTGTTGAGGCTTGAAGTGTGGATTAGAATCAGAAAAAGGTTA 1020
 DB 961 GAGTTAGATCATTTCTGTTGAGGCTTGAAGTGTGGATTAGAATCAGAAAAAGGTTA 1020
 QY 1021 CAAAGATTTAACAAAAATCAGAGGTTCCAAATGGAATGATGTCAAACCAAGTTGTTG 1080
 DB 1021 CAAAGATTTAACAAAAATCAGAGGTTCCAAATGGAATGATGTCAAACCAAGTTGTTG 1080
 QY 1081 GACCGTTGCAAGAAATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 DB 1081 GACCGTTGCAAGAAATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 QY 1141 GATGCAATAGCTGTTATAGTGTGGAATAATCAGTGGGAAATTTTAAACAGAAAACTTT 1200
 DB 1141 GATGCAATAGCTGTTATAGTGTGGAATAATCAGTGGGAAATTTTAAACAGAAAACTTT 1200
 QY 1201 GAAATTCAGATTTATTTTATATATGTTTAACTGGGTTGATTTGGAAGAGCAACGAAAGT 1260
 DB 1201 GAAATTCAGATTTATTTTATATATGTTTAACTGGGTTGATTTGGAAGAGCAACGAAAGT 1260
 QY 1261 GTACTTGAAGAAAAACGAGAGACTATATATAGGTTTATTA 1299
 DB 1261 GTACTTGAAGAAAAACGAGAGACTATATATAGGTTTATTA 1299

RESULT 2
 AX087878 1763 bp DNA linear PAT 17-MAR-2001
 LOCUS
 DEFINITION
 Sequence 5 from Patent WO0114533.
 AX087878
 VERSION
 AX087878.1 GI:13396871
 KEYWORDS
 SOURCE
 ORGANISM
 Candida albicans
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; mitosporic Saccharomycetales; Candida.
 REFERENCE
 1 Rosemond, J.D. and Schnell, N.F.
 Phosphomethyltransferase Kinase (pmk) gene (ex8) from candida albicans
 Patent: WO 0114533-A 5 01-MAR-2001;
 Astrazeneca AB (SE)
 FEATURES
 source
 Location/Qualifiers
 1..1763
 /organism="Candida albicans"
 /mol_type="unassigned DNA"
 /db_xref="taxon:5476"
 ORIGIN
 Query Match 100.0%; Score 1299; DB 6; Length 1763;
 Best Local Similarity 100.0%; Pred. No. 1.5e-280; Indels 0; Gaps 0;
 Matches 1299; Conservative 0; Mismatches 0;
 QY 1 ATGTCAAAAGCATTTAGTGACCTGGAAGACATTTCTTGCTGGATATTTGGTCTT 60
 DB 89 ATGTCAAAAGCATTTAGTGACCTGGAAGACATTTCTTGCTGGATATTTGGTCTT 148
 QY 61 GAGCAATTTATGATGCTTATGTCAGCATTTGTCATCAGAAATGATGCAATATATACA 120
 DB 149 GAGCAATTTATGATGCTTATGTCAGCATTTGTCATCAGAAATGATGCAATATATACA 208
 QY 121 CCAAAAGAACAGTTTGAAGAAATCTAGATCAAAATTTCTTCAACCCCAATTTGCAAC 180
 DB 209 CCAAAAGAACAGTTTGAAGAAATCTAGATCAAAATTTCTTCAACCCCAATTTGCAAC 268
 QY 181 GGAAATGAGGAATTCATCATATTCATCAATATCAGAGAACCCAGAGAGTTCAAGTCAGC 240
 DB 269 GGAAATGAGGAATTCATCATATTCATCAATATCAGAGAACCCAGAGAGTTCAAGTCAGC 328
 QY 241 ATAAATCCATTTTATAGAGCACTATATTCATGTTTATAGCTTATATTTCAACCGACCGAA 300
 DB 329 ATAAATCCATTTTATAGAGCACTATATTCATGTTTATAGCTTATATTTCAACCGACCGAA 388
 QY 301 GCATTTATCTTGAATATCATCATTTACTCAGACCTTGGAATATCATTCACAAGAAATCT 360
 DB 389 GCATTTATCTTGAATATCATCATTTACTCAGACCTTGGAATATCATTCACAAGAAATCT 448
 QY 361 GAAACCAAGACATCTCGAATGAGGAAACAACTTCTTACCATTCCTGCTGCATTTACC 420
 DB 449 GAAACCAAGACATCTCGAATGAGGAAACAACTTCTTACCATTCCTGCTGCATTTACC 508
 QY 421 GAAATGAAAAAGACCGGATTAGGTTCAATCGGACAGATTAAGTCAAGTTGTTGCCACAAGT 480
 DB 509 GAAATGAAAAAGACCGGATTAGGTTCAATCGGACAGATTAAGTCAAGTTGTTGCCACAAGT 568
 QY 481 TTATTTATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGGCAACGTT 540
 DB 569 TTATTTATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGGCAACGTT 628
 QY 541 GCACAGATTGACATTTGTTATGCCCAAAAAAGATAGGATCTGGGTTGATGTTGCAACT 600
 DB 629 GCACAGATTGACATTTGTTATGCCCAAAAAAGATAGGATCTGGGTTGATGTTGCAACT 688
 QY 601 GCAATTTATGCTGATGTTGATATAGAAAGATTTCAAGCTTGTGATTAATGACGTGTT 660
 DB 689 GCAATTTATGCTGATGTTGATATAGAAAGATTTCAAGCTTGTGATTAATGACGTGTT 748
 QY 661 CAGGTTCTAGAAAGATGATCCTGAGAGTTCCCAAGAGTTGAAAAAATTTGATGAAAGT 720


```

Db      749  CAGGTTCTAGAAAGCATCTGAGAAAGTCCCAAGAGTTGAAAAATTTGATTGAAAGT 808
Qy      721  AACTGGGAATTCAAACATGAAGATGATACCATAGGAAATCAAGTTATTAAGGCT 780
Db      809  AACTGGGAATTCAAACATGAAGATGATACCATAGGAAATCAAGTTATTAAGGCT 868
Qy      781  GACGTCAGAGGTGCTCAGAAAACCCAAATTTGATCAAGATCTCCAAATGAAAAAG 840
Db      869  GACGTCAGAGGTGCTCAGAAAACCCAAATTTGATCAAGATCTCCAAATGAAAAAG 928
Qy      841  GAAAAAGCCGAAAGAAAGCTCTGTTGTATGACCACTTAATATGTCCTTACAGTTT 900
Db      929  GAAAAAGCCGAAAGAAAGCTCTGTTGTATGACCACTTAATATGTCCTTACAGTTT 988
Qy      901  ATGAAGGAATGAGGAAATGCGTGAATAATGCACTCAGACCCAGAGCTTATATTA 960
Db      989  ATGAAGGAATGAGGAAATGCGTGAATAATGCACTCAGACCCAGAGCTTATATTA 1048
Qy      961  GAGTTAGATCATCTGTTGAGCCTTGTGACTGTTGCGATTAAGAAATCAGAAAAAGGTTA 1020
Db      1049  GAGTTAGATCATCTGTTGAGCCTTGTGACTGTTGCGATTAAGAAATCAGAAAAAGGTTA 1108
Qy      1021  CAGCATTTAACCAAAAATCAGAGGTTCCAAATTGAACCTGATGTCAAACCAAGTTTGG 1080
Db      1109  CAGCATTTAACCAAAAATCAGAGGTTCCAAATTGAACCTGATGTCAAACCAAGTTTGG 1168
Qy      1081  GACCGTTGCAAGAGATTCCTGTTGTGTGTTGGTGGTTCGAGGTCGTGATGATC 1140
Db      1169  GACCGTTGCAAGAGATTCCTGTTGTGTGTTGGTGGTTCGAGGTCGTGATGATC 1228
Qy      1141  GATGCAATAGCTGTATTAAGTGTGAAAAATCAAGTGGAAATTTTAAGCAGAAAACTCTT 1200
Db      1229  GATGCAATAGCTGTATTAAGTGTGAAAAATCAAGTGGAAATTTTAAGCAGAAAACTCTT 1288
Qy      1201  GAAAAATCCAGATTTATTTTCAATAATGTTTACTGCGTTGATTTGGAAGAGCAAAAGAGGT 1260
Db      1289  GAAAAATCCAGATTTATTTTCAATAATGTTTACTGCGTTGATTTGGAAGAGCAAAAGAGGT 1348
Qy      1261  GTACTTGAAGAAAAACCGAAGACTATATAGTTTATTA 1299
Db      1349  GTACTTGAAGAAAAACCGAAGACTATATATAGTTTATTA 1387

```

```

RESULT 3
LOCUS   AX240804      1299 bp      DNA      linear      PAT 26-SEP-2001
DEFINITION
Sequence 42 from Patent WO0160975.
ACCESSION
AX240804
VERSION
AX240804.1 GI:15797740
KEYWORDS
SOURCE
ORGANISM
Candida albicans
Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE
1 Roemer T., Jiang B., Boone C. and Bussey H.
Gene disruption methodologies for drug target discovery
Patent: WO 0160975-A 42 23-AUG-2001;
JOURNAL
Elitra Pharmaceuticals, Inc. (US)
FEATURES
source
1. 1299
/organism="Candida albicans"
/mol_type="unassigned DNA"
/db_xref="taxon:5476"

```

ORIGIN

```

Query Match      99.0%; Score 1286.2; DB 6; Length 1299;
Beet Local Similarity 99.4%; Pred. No. 1,1e-277;
Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
1 ATGTCAAAAAGCATTTAGTGCACCTGGAAGAAAGCATTTCTTGCTGTGATATTTGTTCTT 60
|||||

```

```

Db      1  ATGTCAAAAAGCATTTAGTGCACCTGGAAGAAAGCATTTCTTGCTGTGATATTTGTTCTT 60
Qy      61  GAGCCAAATTTATGATGCTTATGTCAGACATGTCATCAAGATCATGCACTTATTA 120
Db      61  GAGCCAAATTTATGATGCTTATGTCAGACATGTCATCAAGATCATGCACTTATTA 120
Qy      121  CCAAAAAGAACAGTTTGAAGATCTAGATCAAAAATTTCTTCAACCCCAATTTGCAAC 180
Db      121  CCAAAAAGAACAGTTTGAAGATCTAGATCAAAAATTTCTTCAACCCCAATTTGCAAC 180
Qy      181  GGAGAAATGGGAATATCAATATCATCAATATACAGAGAAAGCCAGAGAGTTGACGAC 240
Db      181  GGAGAAATGGGAATATCAATATCATCAATATACAGAGAAAGCCAGAGAGTTGACGAC 240
Qy      241  ATAAATCCATTTTATGAGGCACTATATTCATCGTTTATAGCTTATATTTCAACCGCA 300
Db      241  ATAAATCCATTTTATGAGGCACTATATTCATCGTTTATAGCTTATATTTCAACCGCA 300
Qy      301  GCATTTGATCTTGAATATCATTTTACTCAGACCCCTGATATCATTCACAAGATATCT 360
Db      301  GCATTTGATCTTGAATATCATTTTACTCAGACCCCTGATATCATTCACAAGATATCT 360
Qy      361  GAAACCAAGACATCTCTGAAATGGAAGAAAAATTTCTTTACCATTTCTGTCCTTACC 420
Db      361  GAAACCAAGACATCTCTGAAATGGAAGAAAAATTTCTTTACCATTTCTGTCCTTACC 420
Qy      421  GAATGGAAGAAAGCCGATTTAGTTCATCGGAGGATTAAGTCAAGTTGTCACCAAGT 480
Db      421  GAATGGAAGAAAGCCGATTTAGTTCATCGGAGGATTAAGTCAAGTTGTCACCAAGT 480
Qy      481  TTATTTATCCATTTTATCCCAATGTTATCAGTACGAATAAAGATATTTTGACAACGT 540
Db      481  TTATTTATCCATTTTATCCCAATGTTATCAGTACGAATAAAGATATTTTGACAACGT 540
Qy      541  GCACGATTTGCAATTTGTTATCCCAAAAAAGATAGATCTGCGTTGATGTTGCAACT 600
Db      541  GCACGATTTGCAATTTGTTATCCCAAAAAAGATAGATCTGCGTTGATGTTGCAACT 600
Qy      601  GCAATTTATGCTGATTTGATATAGAAATTTGACGCACTTGTATTAATGACGTTT 660
Db      601  GCAATTTATGCTGATTTGATATAGAAATTTGACGCACTTGTATTAATGACGTTT 660
Qy      661  CAGGTTCTAGAAAGTATCTGAGAAAGTCCCAAGAGTTGAAAAATTTGATTGAAAGT 720
Db      661  CAGGTTCTAGAAAGTATCTGAGAAAGTCCCAAGAGTTGAAAAATTTGATTGAAAGT 720
Qy      721  AACTGGGAATTCAAACATGAAGATGATACCATACGGAATCAAGTTATTAAGGCT 780
Db      721  AACTGGGAATTCAAACATGAAGATGATACCATACGGAATCAAGTTATTAAGGCT 780
Qy      781  GACGTCAGAGGTGCTCAGAAAACCCAAATTTGATCAAGATCTCCAAATGAAAAAG 840
Db      781  GACGTCAGAGGTGCTCAGAAAACCCAAATTTGATCAAGATCTCCAAATGAAAAAG 840
Qy      841  GAAAAAGCCGAAAGAAAGCTCTGTTGTATGACCACTTAATATGTCCTTACAGTTT 900
Db      841  GAAAAAGCCGAAAGAAAGCTCTGTTGTATGACCACTTAATATGTCCTTACAGTTT 900
Qy      901  ATGAAGGAATGAGGAAATGCGTGAATAATGCACTCAGACCCAGAGCTTATATTA 960
Db      901  ATGAAGGAATGAGGAAATGCGTGAATAATGCACTCAGACCCAGAGCTTATATTA 960
Qy      961  GAGTTAGATCATCTGTTGAGCCTTGTGACTGTTGCGATTAAGAAATCAGAAAAAGGTTA 1020
Db      961  GAGTTAGATCATCTGTTGAGCCTTGTGACTGTTGCGATTAAGAAATCAGAAAAAGGTTA 1020
Qy      1021  CAGCATTTAACCAAAAATCAGAGGTTCCAAATTGAACCTGATGTCAAACCAAGTTTGG 1080
Db      1021  CAGCATTTAACCAAAAATCAGAGGTTCCAAATTGAACCTGATGTCAAACCAAGTTTGG 1080
Qy      1081  GACCGTTGCAAGAGATTCCTGTTGTGTGTTGGTGGTTCGAGGTCGTGATGATC 1140
Db      1081  GACCGTTGCAAGAGATTCCTGTTGTGTGTTGGTGGTTCGAGGTCGTGATGATC 1140

```

QY 1141 GATCAATAGCTGATTAGTTGGTGAATAATCAAGTGGAAATTTTAAAGACAAACTCTT 1200
Db 1141 GATCAATAGCTGATTAGTTGGTGAATAATCAAGTGGAAATTTTAAAGACAAACTCTT 1200
QY 1201 GAAATTCAGATTATTTTCAATTAATGTTTACTGGGTTGATTGGAGAGCAACAGAGGT 1260
Db 1201 GAAATTCAGATTATTTTCAATTAATGTTTACTGGGTTGATTGGAGAGCAACAGAGGT 1260
QY 1261 GTACTTGAAGAAAACCAAGAGCTATATAGTTTATTA 1299
Db 1261 GTACTTGAAGAAAACCAAGAGCTATATAGTTTATTA 1299

RESULT 4
AX488723
LOCUS AX488723 1299 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 6023 from Patent WO02053728.
ACCESSION AX488723
VERSION AX488723.1 GI:22322735
KEYWORDS
SOURCE
ORGANISM
Candida albicans
Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE
AUTHORS Roemer, T., Jiang, B., Boone, C., Bussey, H., and Ohlsen, K.L.
TITLE Gene disruption methodologies for drug target discovery
JOURNAL Patent: WO 02053728-A 6023 11-JUL-2002;
Pharmaceuticals, Inc. (US)
FEATURES
source
1. 1299
Location/Qualifiers
/organism="Candida albicans"
/mol_type="unassigned DNA"
/db_xref="taxon:5476"

ORIGIN

Query Match 99.0%; Score 1286.2; DB 6; Length 1299;
Best local Similarity 99.4%; Pred. No. 1,1e-277;
Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTTTGTGACCTGGAAAAGCATTTTCTGCTGGATATTTGGTCTT 60
Db 1 ATGTCAAAAGCATTTTGTGACCTGGAAAAGCATTTTCTGCTGGATATTTGGTCTT 60
QY 61 GAGCCAAATTTATGATGCTTATGTCAGAGATTTGATCAACGAAATGACGATTATTA 120
Db 61 GAGCCAAATTTATGATGCTTATGTCAGAGATTTGATCAACGAAATGACGATTATTA 120
QY 121 CCAAAAAGAACCGTTTGAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGCAAC 180
Db 121 CCAAAAAGAACCGTTTGAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGCAAC 180
QY 181 GGAAGATGGGAATATCATATCATCAATACAGAGAGCCCAAGAAATTCAGTCACG 240
Db 181 GGAAGATGGGAATATCATATCATCAATACAGAGAGCCCAAGAAATTCAGTCACG 240
QY 241 ATAAATCCATTTTGAAGCAATATATTCATGCTTTTATGCTTATATTCACCGAGAA 300
Db 241 ATAAATCCATTTTGAAGCAATATATTCATGCTTTTATGCTTATATTCACCGAGAA 300
QY 301 GCATTGATCTTGAATCATATTAATCAAGCCCTGATATTCATTCAAGAAATATCT 360
Db 301 GCATTGATCTTGAATCATATTAATCAAGCCCTGATATTCATTCAAGAAATATCT 360
QY 361 GAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTCATGCTCTGTCATTAAC 420
Db 361 GAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTCATGCTCTGTCATTAAC 420
QY 421 GAACTGAGAAAGACCGGATAGGTTTATGGGAGAGATTAATGTCAGTTGTCACAACT 480
Db 421 GAACTGAGAAAGACCGGATAGGTTTATGGGAGAGATTAATGTCAGTTGTCACAACT 480

QY 481 TTATATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGCACAGCTT 540
Db 481 TTATATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGCACAGCTT 540
QY 541 GCACAGATTTGACATTTGATATCCCAAGAAATGATAGATCTGGTTTATGTCACACT 600
Db 541 GCACAGATTTGACATTTGATATCCCAAGAAATGATAGATCTGGTTTATGTCACACT 600
QY 601 GCAATTTATGCTGATGTTATATAGAAATTTGACGAGCTTTGATTAATGAGCTGTT 660
Db 601 GCAATTTATGCTGATGTTATATAGAAATTTGACGAGCTTTGATTAATGAGCTGTT 660
QY 661 CAGGTTCTGAAAGATGATCTGAGAGCTTCCCAAGATTTGAAATTTGATTAAGT 720
Db 661 CAGGTTCTGAAAGATGATCTGAGAGCTTCCCAAGATTTGAAATTTGATTAAGT 720
QY 721 AACTGGAAATTCAAACATGAAAGATGATACATTAATACGGAATCAAGTTATTAAGCT 780
Db 721 AACTGGAAATTCAAACATGAAAGATGATACATTAATACGGAATCAAGTTATTAAGCT 780
QY 781 GACGTCAAGGGGTGCTGAGAAACCCCAATTTGTTATCAGAGTACTCCAAATGAAAAG 840
Db 781 GACGTCAAGGGGTGCTGAGAAACCCCAATTTGTTATCAGAGTACTCCAAATGAAAAG 840
QY 841 GAAAGCCAGAAAGAAAGCTCTGTTGTTATGACAGAGTTAATGCAATTTACGTTT 900
Db 841 GAAAGCCAGAAAGAAAGCTCTGTTGTTATGACAGAGTTAATGCAATTTACGTTT 900
QY 901 ATGAAGGATTTGAGGGGAAATGCGTGAATAATACGATCAGACCCAGAGACTTATTA 960
Db 901 ATGAAGGATTTGAGGGGAAATGCGTGAATAATACGATCAGACCCAGAGACTTATTA 960
QY 961 GAGTTAGATCATTTCTGTTAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAG 1020
Db 961 GAGTTAGATCATTTCTGTTAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAG 1020
QY 1021 CAAGCATTTAACAAATAATCAAGAGTTCCATTTGAACCTGATGTCAAACCCAGTTT 1080
Db 1021 CAAGCATTTAACAAATAATCAAGAGTTCCATTTGAACCTGATGTCAAACCCAGTTT 1080
QY 1081 GACCGTTTCAAGAGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
Db 1081 GACCGTTTCAAGAGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
QY 1141 GATGCAATGCTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
Db 1141 GATGCAATGCTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
QY 1201 GAAATTCAGATTTATTTTCAATATGTTTACTGGGTTGTTGTTGTTGTTGTTG 1260
Db 1201 GAAATTCAGATTTATTTTCAATATGTTTACTGGGTTGTTGTTGTTGTTGTTG 1260
QY 1261 GTACTTGAAGAAAACCAAGAGCTATATAGTTTATTA 1299
Db 1261 GTACTTGAAGAAAACCAAGAGCTATATAGTTTATTA 1299

RESULT 5

AX087876 577 bp DNA linear PAT 17-MAR-2001
LOCUS AX087876
DEFINITION Sequence 3 from Patent WO0114533.
ACCESSION AX087876
VERSION AX087876.1 GI:13396869
KEYWORDS
SOURCE
ORGANISM

Candida albicans
Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE
AUTHORS Rosamond, J.D. and Schneil, N.F.
TITLE Phosphomevalonate kinase (pmk) gene (erg8) from candida albicans
JOURNAL Patent: WO 0114533-A 3 01-MAR-2001;
Astrazeneca AB (SE)

FEATURES

Location/Qualifiers
1..577
/organism="Candida albicans"
/mol_type="unassigned DNA"
/db_xref="taxon:5476"

ORIGIN

Query Match 40.1%; Score 521; DB 6; Length 577;
Best Local Similarity 97.1%; Pred. No. 3.5e-106;
Matches 561; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

177 AACCGAGAGTGGAAATACATATCATCAATACAGAGAACCCAGAGAGTTCAGTC 236
177 AACCGAGAGTGGAAATACATATCATCAATACAGAGAACCCAGAGAGTTCAGTC 236
577 AACCGAGAGTGGAAATACATATCATCAATACAGAGAACCCAGAGAGTTCAGTC 519
237 ACGCAATTAAT-CCATTTTAAAGGCAACATATTCATG- -TTTAAAGTTATATCAACC 293
518 ACGCAATTAATCCATTTTAAAGGCAACATATTCATGCGGTTTAAAGTTATATCAACC 459
294 GACCGAGACATTTGATCTTGAATCATCTTACTCAGACCTCGATCATTCACAGA 353
458 GACCGAGACATTTGATCTTGAATCATCTTACTCAGACCTCGATCATTCACAGA 399
354 AGATACGAAACCAAGACATCTCGAATGAGAGAAAAACATTTCTTACCATTCGTCG 413
398 AGATACGAAACCAAGACATCTCGAATGAGAGAAAAACATTTCTTACCATTCGTCG 339
414 CATTCACGAGTGAAGAGAACCGGATTTAGGTCATCGGAGATTAAGTTCAGTTGGC 473
338 CATTCACGAGTGAAGAGAACCGGATTTAGGTCATCGGAGATTAAGTTCAGTTGGC 279
474 CACAAGTTATATCCATTTTATCCCAATGTATACAGTACGAATTAAGATATTTTGA 533
278 CACAAGTTATATCCATTTTATCCCAATGTATACAGTACGAATTAAGATATTTTGA 219
534 CAACGTTGCAAGATTCGACATTTGTTATGCCCAAAAAAAGATGAGATCTGGGTTGATG 593
218 CAACGTTGCAAGATTCGACATTTGTTATGCCCAAAAAAAGATGAGATCTGGGTTGATG 159
594 TGCACCTGCAATTTATGCTGCTGATGATATAGAAAGATTTTCAGCCAGCTTTGATTAATGA 653
158 TGCACCTGCAATTTATGCTGCTGATGATATAGAAAGATTTTCAGCCAGCTTTGATTAATGA 99
654 CGTGTTCAGGTTCTAGAAAGTATCCTGAGAAAGTCCCAAGAGTTGAAAAAATTTGAT 713
98 CGTGTTCAGGTTCTAGAAAGTATCCTGAGAAAGTCCCAAGAGTTGAAAAAATTTGAT 39
714 TGAAGTAACTGGGAATTCAAACATGAAAGATGTAAT 751
38 TGCAGTAACTGGGAATTCAAACATGAAAGATGTAAT 1

RESULT 6
AX087874 547 bp DNA linear PAT 17-MAR-2001
LOCUS AX087874 Sequence 1 from Patent WO0114533.
DEFINITION AX087874
ACCESSION AX087874
VERSION AX087874.1 GI:13396867

KEYWORDS
ORGANISM
SOURCE
Candida albicans
Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE
AUTHORS
TITLE
JOURNAL
1 Rosemond, J.D. and Schnell, N.F.
Phosphomevalonate kinase (pmk) gene (erg8) from candida albicans
Patent: WO 0114533-A 1 01-MAR-2001;
Astrazeneca AB (SE)
Location/Qualifiers

FEATURES
source
1..547
/organism="Candida albicans"
/mol_type="unassigned DNA"
/db_xref="taxon:5476"

ORIGIN

Query Match 36.2%; Score 470.4; DB 6; Length 547;
Best Local Similarity 99.8%; Pred. No. 7.6e-95;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

828 CCAATGAAAAAGAAAAAGCAAGAAAGCTCTGTTGTATGACCAAGCTTAATAGTGC 887
1 CCAATGAAAAAGAAAAAGCAAGAAAGCTCTGTTGTATGACCAAGCTTAATAGTGC 60
888 CAATTAACAGTTAATGAAGAAATGAGGAAATGCGTAAAAATCGACTGACCCCGA 947
61 CAATTAACAGTTAATGAAGAAATGAGGAAATGCGTAAAAATCGACTGACCCCGA 120
948 GACTATATTAATGAAGAGTATGATCATCTGTTGAGCCTTGAAGCTTGAAGAAACAT 1007
121 GACTATATTAATGAAGAGTATGATCATCTGTTGAGCCTTGAAGCTTGAAGAAACAT 180
1008 CAGAAAGAGTTTACAGACATTAACAACAAATCAGAGGTTCCATTTGAACCTGATGTCA 1067
181 CAGAAAGAGTTTACAGACATTAACAACAAATCAGAGGTTCCATTTGAACCTGATGTCA 240
1068 AACCCAGTTGTTGACCGTTGTCAGAGATTCCTGTTGTGTTGTTGTTGTTGTTGTTG 1127
241 AACCCAGTTGTTGACCGTTGTCAGAGATTCCTGTTGTGTTGTTGTTGTTGTTGTTG 300
1128 TGCTGTGATPACATGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1187
301 TGCTGTGATPACATGACATGATGATGATGATGATGATGATGATGATGATGATGATG 360
1188 GCAGAAACCTCTTGAATTCAGATTAATTTTCAATATGTTTACTGCGTTGATTTGAGCA 1247
361 GCAGAAACCTCTTGAATTCAGATTAATTTTCAATATGTTTACTGCGTTGATTTGAGCA 420
1248 GCAGAAACCTCTTGAATTCAGATTAATTTTCAATATGTTTACTGCGTTGATTTGAGCA 1299
421 GCAGAAACCTCTTGAATTCAGATTAATTTTCAATATGTTTACTGCGTTGATTTGAGCA 472

RESULT 7
CR382139_14/c
wPcOMMENT
Sequence split into 21 fragments LOCUS CR382139 Accession CR382139

Fragment Name	Begin	End
CR382139_00	1	110000
CR382139_01	100001	210000
CR382139_02	200001	310000
CR382139_03	300001	410000
CR382139_04	400001	510000
CR382139_05	500001	610000
CR382139_06	600001	710000
CR382139_07	700001	810000
CR382139_08	800001	910000
CR382139_09	900001	1010000
CR382139_10	1000001	1110000
CR382139_11	1100001	1210000
CR382139_12	1200001	1310000
CR382139_13	1300001	1410000
CR382139_14	1400001	1510000
CR382139_15	1500001	1610000
CR382139_16	1600001	1710000
CR382139_17	1700001	1810000
CR382139_18	1800001	1910000
CR382139_19	1900001	2010000
CR382139_20	2000001	2051428

Continuation (15 of 21) of CR382139 from base 1400001 (CR382139 Debaryomyces hansenii c1)

Query Match 24.9%; Score 323.6; DB 8; Length 110000;
Best Local Similarity 56.1%; Pred. No. 3.7e-62;
Matches 778; Conservative 0; Mismatches 514; Indels 96; Gaps 5;
5 CAAAAGCATTTAGTSCACCTGGAAAAAGATTTCTTGCTGTGATATTTGTTGTTGAC 64

```

Db      87295 CTAGAGATTTCAGTCTCTTGGAAGAAATTTTGGGGGTGGATACCTCGATTAGAAC 87236
QY      65 CAATTTATGATGCTTATGTGACAGCATTTGTCATCGAATGCAATGATATTAACACCA 124
Db      87235 CAATATTCANCTTATGTGACAGCGCTTTCTTCAGAAATGCAATGATTTGTTAAAAA 87176
QY      125 AAGAACCAAGT---TGAAAGAACTAGATCAAAATTTCTTACCCCAATTTGCAAG 181
Db      87175 GAGATTGAGATGAAATGTCATCGTCAACATTCATTAAGCTCCCCCAATTCAGAGG 87116
QY      182 GAGATGGGAATATTCACATP-----TCATCAAAATCAGAGAAGCCAGAGAGTTGAGT 235
Db      87115 GTGATGGAATACGAGTATGATTAATGAAATTTAGTAAAGATTCCTACAGAGTGAAC 87056
QY      236 CAGGCAATTAATTCATTTTATAGAGCAATATATTCATCGTTTACCTTATATTCACCA 295
Db      87055 AAGCTATATATCGGTTTATAGAGGCTACGGTTTACGTTTTCATATATTCACCA 86996
QY      296 CCGAGCATTTGATCTTGAATATCATCTTACTGACCCCTGGATATTCACAGAG 355
Db      86995 AGGAAAAATTCATCTTGAATATCAATTTTTCAGACCCAGGCTATCATCAACATG 86936
QY      356 ATACTGAATCCAGACATCTCGAATGAGAAAAACATTTCTTACCATCTGTCGA 415
Db      86935 ACACGACGCTAAATATCGTCAACAGACAAAGAAATTTTGTACCATTCGAACCA 86876
QY      416 TTACCGAAGTGGAAAAAGCCGATTAAGTTTCACTGCGAGATTAGTCAAGTTGTCGA 475
Db      86875 TTAATGCGGTGCGCAAAACGTATATAGTTCTCTGCTGGGTGGTACAGTATCACTA 86816
QY      476 CAAGTTTATATCCCATTTTATCCCAATGTTATCAGTACGAATTAAGATTTTGCACA 535
Db      86815 CCGGTTGATATTCGCTTTTGAAGACGGTACTATAGTAACTTGCAAAATGATACACA 86756
QY      536 ACGTTGCACAGATTGCACTGTTATGCCCAGATTAAGATGAGATCTGGGTTGATGTTG 595
Db      86755 ATGCTGCCAATATGCCCACGCTATGCGACGAGAAATATGAGTTTGAATTTG 86695
QY      596 CAATCGAATTTATGCTGATTTGATATGAGATTTTTCAGCAGCTTGTGATTAATGAG 655
Db      86695 CTACAGAGCTATAGGCTCATATATATGAGATTCATCTGCTGTGATTAATGAT 86636
QY      656 TGTTCAGTCTTCAAGAAAGATCT-----GAGAAGTTCCCA 694
Db      86635 TGTTCACACGAGATTTTACTCTCAAAATGATGACAGTTGAAGCAGACTACGCA 86576
QY      695 CAGAGTTGAAAAATGATTAAGTAAGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 754
Db      86575 CTGCTTAAGTCACTTAGTGAAGTCCAAATGGGAATTCATTAACCTAGATGACCTTAC 86516
QY      755 CATACGGAATCAGATTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 814
Db      86515 CTCAGGCAATAGATTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 86456
QY      815 TATCAGAGTACTCAATGAGAAAAAGAAAGCCAGAGAAAGCTCTGTGTGTATAC 874
Db      86455 TGTTCAGAGTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 86396
QY      875 AGCTTAATAGTCAATTTAAGTTTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 924
Db      86395 GTCTAACACAGCAATCTGACCTTCATTAATGAGTGAAGTGAAGTGAAGTGAAGTGA 86336
QY      935 ACTCAGACCCAGAGCTTATATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 973
Db      86335 CTGAACACAGCAACCTACAAAGATTTTCTCGATATTTACAGTCTACAGCGTGC 974
QY      974 -----CTGTGAGCTTTGAGT 991
Db      86275 TACTTGATCAATCTGACGAAACGACCCCAATTTGCCCCATTCGTCATTTGATG 86216
QY      992 TTGGAATTAAGAACTCAGAAAGGTTTACAGATTTAACAACAAATCAGAGTTCCA 1051
Db      86215 AGGCAATCAAAATATATAGATCAATTTGAGATCTAACCGCTTACTGCGGTGAGAA 86156

```

```

QY      1052 TTGAACCTGATGTCACAAACCCAGTTGTGAGCGTTGCAAGAGATTCCTGTTGTG 1111
Db      86155 TAGAACCCCAATCCAGACTATCTTTTGATATCTGATACAGATTAAGGAGTCTG 86096
QY      1112 GTGTGTGTTTCCAGGTGCTGTGTGATGATGATGATGATGATGATGATGATGATG 1171
Db      86095 GTGGATGTTTCCAGGTGCTGTGTGATGATGATGATGATGATGATGATGATGAT 86036
QY      1172 AAGTGGAAATTTTAAACAGAAACCTTGAATAATCCAGATTAATTTTATATAGT 1231
Db      86035 CGATCAATGAATTTGTGATTCAC---CAAAAGACAAATAGTTGAGATGTCGAT 85979
QY      1232 GGTTGATTTTGAAGAGCAACAGAGTGTACTTGAAGAAAAACAGAGACTATATAG 1291
Db      85978 GGTGATTTTGAAGAGCAACAGAGTGTACTTGAAGAAAAACAGAGACTATATAG 85919
QY      1292 GTTTATTA 1299
Db      85918 GCTTATTA 85911

```

RESULT 8 CR380952.03 WPCOMMENT

Sequence split into 10 fragments LOCUS CR380952 Accession CR380952

Fragment Name	Begin	End
CR380952.00	1	110000
CR380952.01	100001	210000
CR380952.02	200001	310000
CR380952.03	300001	410000
CR380952.04	400001	510000
CR380952.05	500001	610000
CR380952.06	600001	710000
CR380952.07	700001	810000
CR380952.08	800001	910000
CR380952.09	900001	927101

Continuation (4 of 10) of CR380952 from base 300001 (CR380952 Candida glabrata strain CB

Query Match 15.5%; Score 201; DB 8; Length 110000;
Best Local Similarity 51.2%; Pred. No. 1,1e-34;

Matches 688; Conservative 0; Mismatches 595; Indels 60; Gaps 7;

```

QY      1 ATGTCAAAAGCATTTAGTGAACCTGGAAGAACATTTCTGCTGTGATATTTGGTCTT 60
Db      90882 ATGGAAGAGCATACAGTCTCTCAAGTAACTTTACTGTTGCTGTGATATTTGG 90941
QY      61 GAGCCATTTATGATCTTATGATGACAGCATTTGATCAGCAATGATGATGATGATG 116
Db      90942 GATCCAAATATGACTCTTACGTTGTAGCTGATCATCGGAATGATGCTGTTGTCG 91001
QY      117 -AACCAAAAGAACCGATTGGAAGATCTGAAATCAAAATTTCTTACCCCAATTTG 175
Db      91002 GAAAGGAAGTATACGAACTTTAGATGCTTCCATTCAGTTTACAGTCTCAATTT 91061
QY      176 CAACGGAATGGAAT---ATCATAATCAATCAATCAGAGAGCCAGAGAGT 231
Db      91062 AATATATGTTGATGATTAATGATTAATTTCAAAATGTTATGTCAGATCGAAGTT 91121
QY      232 CAGTCAAGCATTAATTCATTTTATAGGCAACTATATTCATGTTTATGATTAATTC 291
Db      91122 TCGGGAATGCAATTCATTTTATAGAGAACTGTGTTAAAGTTTATGAGT 91181
QY      292 CCGACGAGCATTTGATCTTGAATATCATCTTACTGACCCCTGGATATCATCA 351
Db      91182 CTGTGATGTTAAACGTTTCCAAATGTTATCAAGATATGATTTACTGAGTCAAGT 91241
QY      352 GAAGATCTGAACCAAGCATTC-----TCGAATGAGAAAAACATTTCTTATCCAT 405
Db      91242 TATCATATGCAAAACATATGATCTATGATTCATATGCTTGAAGAACTTCAAGTT 91301
QY      406 TCTGTGCAATTAACGAGTGAAGAAAGCCGATTAAGTTGATGCGGATTAAGTGTCA 465

```

Db 91302 AAGACCGGTATAGCCCAAGTCCAAAACAGATTGGATCATCTGCTGCTAGTTG 91361
 Qy 466 GTTGTGGCACAAGTTATATCCCAT---TTATCCCAATGTATCAGTACGAATAA 522
 Db 91362 GTCTGTACACATGCATTAATTTGAGTTTCTTAAATGACATTCATGCTCGATGACTT 91421
 Qy 523 GATATTTTGCACACGTTGCACAGATTTGCATTTGTTATGCCCAAAAAAGATGAGATCT 582
 Db 91422 GATTCATCATTAATATCATCAATATATCATATGTCAGGACAGGCAAGATCGGCACT 91481
 Qy 583 GGGTTGATGTTGCACTGCAATTTATGCTGATTTATATGAAAGATTTTCAAGCCACT 642
 Db 91482 GGGTTGATGTTGCAAGCCGCTGTTATATGTTCAATTAATGATGCAAGATTTTCAACTGAT 91541
 Qy 643 TTGATTAATAGACGTGTTTCAAGTTCTAGAAAGATCCTGAGAAAGTTCCCAAGAGTTG 702
 Db 91542 TTAAATATCTAGACTACTTCTTATATAGTATTAATCTACAGAGATGACAGATTCGAATG 91601
 Qy 703 AAAAAATTGATGAAAGTAACTGGAATTCAAACATGAAAGATGTACATTACCATACGA 762
 Db 91602 AGTAAATTAATAAATCAATGGAATTAATAAATCAATGATGCTGTTGCCAGAGAGG 91661
 Qy 763 ATCAAGTTATTAATGGGTGACGTCAAGGGTGGCTCAAAAACCCCAATTTGATATACGA 822
 Db 91662 TTGAGACTGTGTGGGTGATGTAAGGGTGGTTCAGAGACTGTTGAGTTGGAAAAA 91721
 Qy 823 GTACTCCAAATGAAAAAGAAAAAGCAAGAAAGCTGTTGTATGACCAAGCTTAAT 882
 Db 91722 GTGAAATTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 91781
 Qy 883 AGTGCATTTTACAGTTTATGAAAGAAATGAGGAAATGCGTAAAAATATACAGCTCAGC 942
 Db 91782 GAGGCAATTTGATGTTTATAGATGCGATGCTTCAATTAAGCAATTAACAAAGCTGAT 91841
 Qy 943 CCAAGACTTAATA-----TTAAAGATTGATCATTCTGTGA----- 980
 Db 91842 AATGACAAATGCAAAAGTCTCTTGAAGGCAATTAATTAATTAATTAATTAATTAAT 91901
 Qy 981 -----GCTTTGACCTGTGGATTAAGAAATAGAAAAAGGTTACAGATTAAGA 1032
 Db 91902 GAAATATCTTCACTGAAACCGGCAATTTGCAAAATTAAGGAAATTTCAAGCTAATTAAG 91961
 Qy 1033 CAAAAATCAGAGGTTCCAAATTTGAACCTGATGTCCAAAACCAAGTTGTCGCGTTGCA 1092
 Db 91962 CAAAGTCTGATGCTGATTAATTAAGCAAGGCTCAAAACCAATTAATTAAGTCTGTTCA 92021
 Qy 1093 GAGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1152
 Db 92022 CAGTTAAATGTTGTTATTTGAGGAGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 92081
 Qy 1153 GATTAAGTGTGGAATAAATCAAGTGGAAATTTTAAGCAGAAAACTTTGAAATTCAGAT 1212
 Db 92082 ATTAATCTCAAGAAATACAAATTTGAAATGAAAAACAATGCTCAGAA----- 92132
 Qy 1213 TATTTTCATATGTTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1272
 Db 92133 ---TTGATTCGCTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 92189
 Qy 1273 AAACCAAGAACTATATAGGTTT 1295
 Db 92190 AACCTGAATATTAACAGAGATTT 92212

RESULT 9
 AE016898 314422 bp DNA linear JUN 14-APR-2004
 LOCUS AE016898 Ereothecium gossypii chromosome V, section 5 of 5 of the complete
 DEFINITION sequence.
 ACCESSION AE016898 AE016818
 VERSION AE016898.1 GI:44983975
 KEYWORDS Ereothecium gossypii (Ashbya gossypii)
 SOURCE Ereothecium gossypii
 ORGANISM Ereothecium gossypii

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 15001715
 2 (bases 1 to 314422)
 Gates, K., Dieckhoff, F. S., Brachet, S., Voegel, S. E., Lerch, A.,
 Philippsen, P., and Galfrey, T.
 Direct Subdivision
 Submitted (20-DEC-2002) Department of Ashbya Genomics, Syngenta,
 Research Triangle Park, NC 27709, USA
 This is low coverage sequence generated to identify the complete
 set of genes and the gene order on this chromosome. Regions of low
 quality are identified. Before doing extensive work on any gene
 identified here the sequence should be confirmed.

FEATURES
 source
 1. 314422
 /organism="Ereothecium gossypii"
 /mol_type="genomic DNA"
 /strain="ATCC 10895"
 /db_xref="ATCC:10895"
 /db_xref="taxon:33169"
 /chromosome="v"
 /note="synonym: Ashbya gossypii"
 1. 33050
 /organism="Ereothecium gossypii"
 /mol_type="genomic DNA"
 /db_xref="taxon:33169"
 /clone="BAC clone BAG1515"
 16499. 70444
 /organism="Ereothecium gossypii"
 /mol_type="genomic DNA"
 /db_xref="taxon:33169"
 /clone="BAC clone BAG1624"
 34420. 103446
 /organism="Ereothecium gossypii"
 /mol_type="genomic DNA"
 /db_xref="taxon:33169"
 /clone="BAC clone BAG1078"
 108821. 151176
 /organism="Ereothecium gossypii"
 /mol_type="genomic DNA"
 /db_xref="taxon:33169"
 /clone="BAC clone BAG1191"
 154249. 227554
 /organism="Ereothecium gossypii"
 /mol_type="genomic DNA"
 /db_xref="taxon:33169"
 /clone="BAC clone BAG1362"
 204900. 270712
 /organism="Ereothecium gossypii"
 /mol_type="genomic DNA"
 /db_xref="taxon:33169"
 /clone="BAC clone BAG1577"
 244822. 312407
 /organism="Ereothecium gossypii"
 /mol_type="genomic DNA"
 /db_xref="taxon:33169"
 /clone="BAC clone BAG1474"
 93
 /note="region of low quality - less than 90 percent
 confidence"
 complement(1..252. >1559)
 /locus_tag="AER306C"
 complement(1..252. >1559)
 /locus_tag="AER306C"
 /product="AER306Cp"
 complement(252. 1559)

unsure
 gene
 mRNA
 CDS

/locus_tag="AER306C"
/note="syntenic homolog of Saccharomyces cerevisiae
YML032C (RAD52) "
/codon_start=1
/evidence=not experimental
/product="AER306Cp"
/protein_id="AAS52986.1"
/db_xref="GI:44983976"
/translation="MADDKKQGVVDDIOAKLDRIGLPEYISRVGHGTRVAYIES
WKAINLANQIFGPNWGSSEKSVTVDFMDERQGRPTIGCTYDQPLSKIDYKFEEDIG
FDEGNLFRAIDENNELSRANTVYTDSEGFARKNVAPLPAPRGAMRPEAAHAPAP
WBPQAPVAFVYAKAPDLOAHGVPVPHLLDPKQAOQSIHRTVDQTSAAVKSIVR
EKGIQDARDITSKAPKGIPLPPDITADDPDLAARPAAPONLPTNNSIQRT
MLEPLHFPQPTAVPLPNVPASRREVRPKQSPSLAKPPQ"
<2084. .>3859
/locus_tag="AER307W"
<2084. .>3859
/locus_tag="AER307W"
/product="AER307Wp"
2084. .3859
/locus_tag="AER307W"
/note="syntenic homolog of Saccharomyces cerevisiae
YML031W (NDC1) "
/codon_start=1
/evidence=not experimental
/product="AER307Wp"
/protein_id="AAS52987.1"
/db_xref="GI:44983977"
/translation="MYDSATPAPNLRLSYNALFSDICKRPNVYATRLLLTSLVSG
AAVTIMGRMDINTWERMILPPVVALVYFALLVITTRKYLHVQFLGSSNTTVYG
EFSIKFVLVQITHVGLLISLISGLVYDSCGSIPLVIRLNEMLVPLLYTVQGL
AFDIDRISFVNGHROSQAFITISPRDSFVCCVLAATFVLQPVVYMLTTLKLSQ
FLAYIOTTLAFTIQLMDTITAFNAVHISIGLHGRKIPSLISYPMETLVYGNKSK
RMFRLRAPOELSVRATSEVQWRLEPIYHTRXGNSHITWNIILRECLVITIOENKIAD
FISKIBSKATFSSAKOMFSLSDVDVFGNQPALSTAAVPOAHLPTPSORLNDPL
DSGAGVAAVYFPQSGAROFIQPHASVNPPIFPOQSMHLVDVNAVYTKLVSPFP
PGQANASGRPLSFELMTISKHOAEKLIPLOQSVACVIALMGMLNSIDEPKG
SVYSVGEVLTILERSVGLGTVEWELRSSSDSGDAISVYIDLINVEITLVKN
ELINDVYDDVDVKLSQWVLQTCNE"
2311. .2312
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
2403
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
2405. .2409
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
2421. .2422
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
2433. .2434
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
2437. .2438
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
2454. .2456
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
2825
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"

unSURE 2870. .2872
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
unSURE 2981. .2982
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
unSURE 3527
/locus_tag="AER307W"
/note="region of low quality - less than 90 percent
confidence"
gene complement(<4007. .>5845)
/locus_tag="AER308C"
complement(<4007. .>5845)
/locus_tag="AER308C"
/product="AER308Cp"
complement(4007. .5845)
/locus_tag="AER308C"
/note="syntenic homolog of Saccharomyces cerevisiae
YDR456W (NHX1) "
/codon_start=1
/evidence=not experimental
/product="AER308Cp"
/protein_id="AAS52988.1"
/db_xref="GI:44983978"
/translation="MMANRAIOYMIWTLRMGAAGNDAGGVNEPQGPIDGP
DENVPIITBEIFSSMALCILFLILALMTNRYLNRORRAVHETVLSIFGMUGFL
IRVSPGHYIODTVFNSYFENFLPPIILNSXELNQNTPNSIVLSIFTAIPGFI
SAVLYGILYLTALGIESIDISPLDASVGTALSDVDFLTSIPNAKVDPKLYTI
ILKHSVRDLSIVIFETQCKPRGKLTIFMSFPEGISGLPELMTFSLIGVWGLIVA
RTOGLTIKYIFQVILASENFIIFYGLALFQVEILVYKMLIIVAISSICVACAVF
PLSDFVNWILYRWVWTLKRSKAVSIPIDEIPYOQWMPFACLRGAVGVALMGLOE
FKRPLATVYLWVWLVIIIFGGTAGMLIILNIKTGCVREORDETDEPDIAPRVOL
MSLDADPPPLASPTSGPQSEFPPMIDRPSRSASGADQPVGRSLDRNRITFLAN
<6260. .>6817
/locus_tag="AER309W"
<6260. .>6817
/locus_tag="AER309W"
/product="AER309Wp"
6260. .6817
/locus_tag="AER309W"
/note="syntenic homolog of Saccharomyces cerevisiae
YDR454C (GUK1) "
CDS
Query Match 12.2%; Score 158.8; DB 8; Length 314422;
Best Local Similarity 50.7%; Pred. No. 2.8e-25;
Matches 497; Conservative 0; Mismatches 462; Indels 21; Gaps 4;
QY 1 ATGTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTGCTGGATATTGGTCTT 60
DB 85664 AAGGTAAAGGCTTTAGAGGCCCGCAGTAAGGCTCTTCTAGTTGGGGCTVATTTGGTCTT 60
QY 61 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGCATCAGAAATGCATGAGTTATAC 120
DB 85724 GACCCCAACTACAGTCTTATGTTGTAGCACTTCTCTGCGATGATGAGTAGTAAAG 120
QY 121 C---CAAAAGAACCAAGTTGAAGAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGCA 177
DB 85784 CAGACTGAAGGCTTCGACAGAGGCCCTTACGTCAAGTGAACAAGTTTCAATTTAAAC 177
QY 178 AACGAGAAATGGGAATATCATATCATCAAA--TACAGAGAAAGCCAGAGAAAGTTGAC 234
DB 85844 AACATCAGTGGGAATACACAGTGAAGTTGAGATGCAATTTATTAATTTGTTGACGAAA 234
QY 235 TCACGCAATTAATCATTTTAAAGGCAATATATTTATGCTTTAGCTTATATTTCAACCG 294
DB 85904 GAGCGCGGCAATCATTTATTTAGAAAGACTTACAAATGTAATTTGAATTTTCAACCG 294
QY 295 ACCGAAAGCATTTGATCTTGAATATCATATTTTACTGAGACCCCTGGAAATATTTCAAGAA 354

Db 85964 CCCCCGGGACAGCTTGTGAGATATTCATTTAAAGTTTCTCTGAGCCGGCTAACCATCC 86023
Qy 355 GATACCTGAACCAAGACATCTCGAATGAGAAAAACATTTCTTACCATTTCTCGTCC 414
Db 86024 CAACCTGATAGATTTGAACACCGGATAGTTCGGATAGTTTCAAGATTTTCAAGAAAAAC 86083
Qy 415 ATTACCGAAGTGAAGAAAGCCGATTTAGTTTCACTGGAGATTTAGTGTCACTTTGGCC 474
Db 86084 ATCACCGATGTTTCAAGAGACGGGTGGGGTCACTGACAGGGCTAGTCAACCTTCTCACT 86143
Qy 475 ACAATTTATTTTCCCATTTTATCCCAATG-----TATCGTACGATTAAGT 525
Db 86144 GCGCTCTGATATCTGTGTTTTTACCGAGATGATGTTGCTTAATTAACATTTAGG 86203
Qy 526 ATTTGCAACAGTTGACAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 585
Db 86204 CTGATTCATTAATTTGTCAAGATTCGCGATTTGCAAGCCCAAGAAATAGGCACTGAT 86263
Qy 586 TTTGATGTTGCACTGCAATTTATGTTCTGATTTATATAGAAATTTTCAAGCTTTG 645
Db 86264 TTTGACGTGACCGGATTTTGGCTGATCTTACAGGGCTTTCCGCTGATTTG 86323
Qy 646 ATTAATGACGTGTTTCAAGTTCTAGAAATGATCTGAGAGTTTCCCAAGATTTGAAA 705
Db 86324 ATTAATGATCTGCTCTAATTTG-----GTGGCCGAATATGCAAGTGTTCGCGAAC 86377
Qy 706 AAATGATTAAGATTAATCTGGAATTTCAAAATGAAGATTAATTAATTAATTAATTAATC 765
Db 86378 GTAAATGATTAAGCGGATTTGCGATGCTGCCATGACCGGATGCTGCTCCCAAAAGCTTA 86437
Qy 766 AAGTTATTAATGCTGACGTCAAGGCTGCTCAAGAAACCCAAATTTGATCAGAGTA 825
Db 86438 CGCATCGTATGAGAGATGTCGGCGGGGTTTCAAGAGCGCTGTTGTTCACACAGTG 86497
Qy 826 CTCCAATGAAAAAGAAAAAGCCAGAAAGCTCTGTTGTATGACCAAGCTTAATAGT 885
Db 86498 CACCAATGTTTCAAGCTGATCAGAAACCGTAGCAAGATCTACAGAGAGATTAACAAA 86557
Qy 886 GCCAATTAACAGTTATGAAAGAAATGAGGAAATGCGTGAAGAAATAGACCTCAGCCCA 945
Db 86558 GGAAACATTTGCTTTATGTCGGCTTTGGAAGAACTCAATAGATTTGCAAGACAGATCTT 86617
Qy 946 GAGCTTATTAATTAAGATT 965
Db 86618 AGCACATACATGAACTGCT 86637

RESULT 10
CR382123_06/c

Sequence split into 18 fragments LOCUS CR382123 Accession CR382123

Fragment Name	Begin	End
CR382123_00	1	110000
CR382123_01	100001	210000
CR382123_02	200001	310000
CR382123_03	300001	410000
CR382123_04	400001	510000
CR382123_05	500001	610000
CR382123_06	600001	710000
CR382123_07	700001	810000
CR382123_08	800001	910000
CR382123_09	900001	1010000
CR382123_10	1000001	1110000
CR382123_11	1100001	1210000
CR382123_12	1200001	1310000
CR382123_13	1300001	1410000
CR382123_14	1400001	1510000
CR382123_15	1500001	1610000
CR382123_16	1600001	1710000
CR382123_17	1700001	1753957

Continuation (7 of 18) of CR382123 From base 600001 (CR382123 Kluveromyces laetis strai)

Query Match 11.4%; Score 148.4; DB 8; Length 110000;

Best Local Similarity 49.7%; Pred. No. 6,6e-23;
Matches 593; Conservative 0; Mismatches 541; Indels 60; Gaps 6;

Qy 9 AGCATTTGTCACCTGGAAGACATTTCTGCTGCTGATTTTGGTCTTGACCAT 68
Db 12974 ATCTTTATGCTCCAGGTAAGCCCTTCTACAGCGGTATTTATGTTTAAACCA 12915
Qy 69 TTATGATGCTTATGTCAGCATTTGATTCAGAAATGATGATGATTAATTAACCA 128
Db 12914 ATATGATCTTATGTCGTGGCCCTATCGTCAAGAAATGATGCTGTTGTCACCA 12855
Qy 129 AACAGTTGAAGATC-----TGAATCAAAATTTCTTACCCCAATTTGC 176
Db 12854 ATTAATTTGTCATATGCTCATCAATTAATCTTTAAAGTCAAGATTAATCA 12795
Qy 177 AAACGGAAGTGGAAATTCATATCATCAATTAACAG---AGAACCCAGAAAGTTCA 233
Db 12794 TGACGACAGTGGCAGTATTAAGATGATTCAGCTCAATTAACACCCCAAGAAATGCA 12735
Qy 234 GTACGACATTAATTCATTTTATGAGCACTATATTCATCGTTTATGCTTATTAATCA 293
Db 12734 AGAAAAAGAAATCCCTTTCATGAAAGTTTATCAACAGTCTTCACTTTCAC 12675
Qy 294 GACCGAAGCATTTGATCTTGAATCATCATTTTACAGCCCTGATATCATTTCAACA 353
Db 12674 CGATTTAAACTACAGAGAAACATAGTATGATTTTATCTGAGATGCTGCTATCATTC 12615
Qy 354 AGATATGAAACCAAGACATCTCGAATGAGAAAAAACATTTCTTTACATTTCTGCTGC 413
Db 12614 CCAAGCGGAAACAAATCGAAGATTAAGGATTAAGTAAATCTTTTTCATATGAATATC 12555
Qy 414 CATTAACGAAAGTGAAGAACCGGATTTAGTTTATCGGAGGATTTAGTCTAGTTTGC 473
Db 12554 TATCATGATGCTCCCAAAATCGGTTGGGCTTCAAGCGGATTTGTTTATCTGTTTATG 12495
Qy 474 CACAAATTAATTAATCCATTTTATCCCAATGTTATCAGTATCAATTAAGATTAATTTG-- 531
Db 12494 CACGCAATTTGTTCTGCTTCAACAAATCTTATGTTTATCATCAAGATGATATGAG 12435
Qy 532 -----CACACGTTGACAGATTTGCACTTTGTTATGCCCCAAAAAGATAGATCTGG 584
Db 12434 ATTAATTAACCAATTTGTTCTCAAGTGCATTTGTAGGCAAGGATTAATTAAGTATG 12375
Qy 585 GTTTGATGTCGCACTGCAATTTATGCTGATTTATATATAGAAATTTTACAGCCAGCTT 644
Db 12374 CTTTATGTTGCTGACAGTGTACGAAACATCTTACAAAGATTTGATCCGTACT 12315
Qy 645 GATTAATGACGTGTTTCAAGTTCTTGAAGATGATCTGAGAAATTTCCCAAGATTTGA 704
Db 12314 CATGAGCA-----ATTACATACCAATTTATGTCAGCAATACATTTGCTTTAGTTCT 12260
Qy 705 AAAATGATGAAAGTAACTGGAATTCAAACATGAAGATTTATACATTAAGCAAT 764
Db 12259 TTGATTTGA-TGAACCGAATTTGGAATAAAGGCTCTTGAAGTGTCTTCCCTCAAGATT 12201
Qy 765 CAAGTTATTAATGAGTGAAGTGAAGGAGTGCAGAAACACCAATGTTGTTACGAGT 824
Db 12200 GAGGCTAATTAATGAGGATTTTAAATTAATGATCGAGACACAAATCTGTTGCAAAAGT 12141
Qy 825 ACTCCATGAAAAAGAAAGCCAGAAAGTCTGTTGTGATGTCACAGCTTAATAG 884
Db 12140 CAATGAATGATTAATGAAACGTCAGACGCTTTTAAATTTATTAAGGCATAGACA 12081
Qy 885 TCCCAATTTACAGTTTATGAAGAAATGAGGAAATGCGTGAATAATTAAGCACTCAGACC 944
Db 12080 AGCGAACAAGAAATTTGTCATCTTCGTCGAACTTCTGCTTATGACCAAGCAC 12021
Qy 945 AGAGCATTAAT-----AAAGTTATGATCATTC 974
Db 12020 TTGGAATATGAAGCTTGAATTTGAACATTAATGATTAAGAAAGAAAGACACACCACT 11961
Qy 975 TGTGAGCTTTGACGTGTTGCAATTAAGAAATCAGAAAGGTTTACAAAGCATTAACACA 1034

Db 11960 ACTGAGAAAATGATTCATTCGTTCAATCAATTAAGAAACATTCAGACTCATTACGGA 11901
Oy 1035 AAAATCAGAGGTTCCATTTGAACCTGATGCCAAACCCAGTGTGGACCGTTGTCAGA 1094
Db 11900 ACAATCAGAGTCCGATTTGAACCTGAAAGTCCAAACATCTTTAAATGCTGCAGATCAG 11841
Oy 1095 GATTCTGCTGTGTGTTGGTGTGTGTTCCAGAGTCTGTGTGATTAAGATGCAAT 1148
Db 11840 TTTGCGGGAGTGTAAACCGGAGTTGTACAGAGTCTGAGAGATATGATGCTAT 11787

RESULT 11
YSCER8 2399 bp DNA linear PLN 27-APR-1993
LOCUS YSCER8
DEFINITION S.cerevisiae phosphomevalonate kinase gene, complete cds.
ACCESSION M63648
VERSION M63648.1 GI:553127
KEYWORDS phosphomevalonate kinase.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE Tsay, Y.H. and Robinson, G.M.
1 (bases 1 to 2399)
TITLE Cloning and characterization of ERG8, an essential gene of
Saccharomyces cerevisiae that encodes phosphomevalonate kinase
Mol. Cell. Biol. 11 (2), 620-631 (1991)
JOURNAL 91117228
MEDLINE 1846667
PUBMED
COMMENT On Oct 3, 1994 this sequence version replaced gi:171478.
Original source text: S.cerevisiae DNA.
FEATURES
source location/Qualifiers
1..2399
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
767..2121
/gene="ERG8"
767..774
/gene="ERG8"
847..2121
/gene="ERG8"
/EC_number="2.7.4.21"
/codon_start=1
/product="phosphomevalonate kinase"
/protein_id="AAA34596.1"
/db_xref="GI:171479"
/translation="MSBLRPFSAFGKALLAGYLVDITYEARFVGLSRMNAVHPY
GSLQSGDKFEVVKQFQKDEWLVIHSPKSGFTPVISGSKPTEKVIANVFTEK
PNMDYCNRLPVIDIFSDDAHSOEDSVTEHGNRLSPHSRIEVPKATGSSAG
LVTVLTALASFPVSDLENNVDKYREVIHNLAVOAGCOQKIGSGFDVAAGVGSIR
YRRPPLALISNLPDIGSATVSGKLAIHVEDMDITIKSHLPSGLTMMGDIKNSE
TVKLQVKRWYDSHMPESLKITIYELDHANSRPMQDLSTKIDRLHETHHDYSDQIFESL
ERNDCTQKPEITTEVDAVATIRSFRTKTSAGADIEPPVOTSLIDDCQTLKGYLT
CLIPGAGYDAIVITKQVDVLRQPLMTKD"

gene
TATA_signal
CDS
Query Match 10.2%; Score 132.2; DB 8; Length 2399;
Best Local Similarity 50.6%; Pred. No. 4e-19;
Matches 474; Conservative 0; Mismatches 428; Indels 35; Gaps 5;
Oy 3 GTCAAAAGCAATTTAGTGACCTGGAAAAGCATTTCTTGCTGGTGATATTTGGTCTTGA 62
Db 855 GTTGAGAGCCTTGAGTGCCCAAGGAAAGCGTACTAGCTGGTGATATTTAGTTTGA 914
Oy 63 GCCAATTTATGATCTTATGACAGCATTTGTCTATCAGATGATGCAATTAACACC 122
Db 915 TACAAATATGAAACATTTGATGTCGATTATCGCAGAAATGATGCTG-TACCCATC 973
Oy 123 AAAAGAACAGATTGAAAGAAATCTAGAAATCAAAATTTCTTACCCCAATTTGCAACG 182
Db 974 CTTACGGTTCAATGCAAGGCTGTAGTAAGTTTGAAGTCCGTGCAAAAGTAAACAATTTA 1033

ORIGIN

Query Match 10.2%; Score 132.2; DB 8; Length 2399;
Best Local Similarity 50.6%; Pred. No. 4e-19;
Matches 474; Conservative 0; Mismatches 428; Indels 35; Gaps 5;

Oy 3 GTCAAAAGCAATTTAGTGACCTGGAAAAGCATTTCTTGCTGGTGATATTTGGTCTTGA 62
Db 855 GTTGAGAGCCTTGAGTGCCCAAGGAAAGCGTACTAGCTGGTGATATTTAGTTTGA 914
Oy 63 GCCAATTTATGATCTTATGACAGCATTTGTCTATCAGATGATGCAATTAACACC 122
Db 915 TACAAATATGAAACATTTGATGTCGATTATCGCAGAAATGATGCTG-TACCCATC 973
Oy 123 AAAAGAACAGATTGAAAGAAATCTAGAAATCAAAATTTCTTACCCCAATTTGCAACG 182
Db 974 CTTACGGTTCAATGCAAGGCTGTAGTAAGTTTGAAGTCCGTGCAAAAGTAAACAATTTA 1033

Oy 183 AGAATGGAAATATCACATATCATCAATAACAGAAAG-----CCAGAGAAAGTTT 232
Db 1034 AAGATGGAGAGTGGCGTACATATATAGTCTTAAAGATGGCTTCATTCGTTCCATAG 1093
Oy 233 AGTCAGCATTAATTCATTTTATAGAGCACTATATTCATCGTTATAGCTTATATCAAC 292
Db 1094 GCGGATCTAAGAACCTTTCATTTGAAAGAAAGTTATGCTTAACGATTTAGCTTTAAAC 1153
Oy 293 CGACCGAAGCATTTGATCTTGA-----TCATCATTTATCTGAGACCTGTGATATC 343
Db 1154 CTATCATGACGACACTACTGCAATAGAAACCTGTGCTGATATGATATTTCTGATGATG 1213
Oy 344 ATTCACAAAGATACTGAACCAAGACATCTCGAATGAGAAAGAAACATTTCTTACC 403
Db 1214 CTTACATTTCTAGAGAGATAGCTTATCCGACATCGTGGCAACAGAAATGATGTTTC 1273
Oy 404 ATTTCTGTCATTTACCGAATGGAAGACCGGATAGGTTGATCGGCGAGATATGATG 463
Db 1274 ATTGCAAGAAATGGAAGATTTCCCAAAACAGGCTGGGCTCTCGGCAAGTTTATGTC 1333
Oy 464 CAGTTGTGCGACAAAGTTTATTTATCCCATTTTATCCC-----CAATGTTATCAGTA 514
Db 1334 CAGTTTATCTACAGCTTTGGCCCTCTTTTGTATGAGACCTGGAATAATATGATGACA 1393
Oy 515 CGAATTAAGATATTTGCAACAAGCTGCAAGATGCAATGTTATGATGATGATGATGAT 574
Db 1394 AATATAGAGAAATTTATCATATTTAGCAAGATGTCATATGCAAGCTCAGGATGAAA 1453
Oy 575 TAGATCTGGGTTGATGATGTTGCAACTGCAATTTATGCTGATGTTATATAGAAATTT 634
Db 1454 TTGAAAGCGGTTTATGATGATGAGCGGCGGCAAGATATGATGATGATGATGATGAT 1513
Oy 635 AGCAGCTTTGATTAATGACGCTGTTCAAGTTCTAGAAAGTATCTGAAAGTTCCCA 694
Db 1514 CACCGCATTTATCTCTAATTTGCGAAGATTT-----GGAAGGCTACTTACGCGAGTA 1567
Oy 695 CAGAGTTAAAGAAATTTGATTAAGATTAATCGGGAATCAACATGAAGATGATCATTT 754
Db 1568 AACTGGCGCATTTGGTGTATGAAGAGCTGAAATATTAAGATTAAGATTAACATTTAC 1627
Oy 755 CATACGAATCAAGTTATTAATGAGTGACGTCAGAGGTGCTCAGAAACACCCAAATTTG 814
Db 1628 CTTGGGATTTAATCTTATGATGAGGCGATTTAAGAAATGTTTCAGAAACAGATTAAC 1687
Oy 815 TATCAAGATCTCAATGAGAAAGAAAGAAAGCCAGAAAGATCTGTTGATGATGCC 874
Db 1688 TCCAGAAAGTAAAGAAATTTGATGATGATTTGCAATGCGCAAGAGCTTGAAATATAT 1747
Oy 875 AGCTTAATAGCGCAATTTACAGTTTATGAAGAAATT 911
Db 1748 AACTGATCATGCAAAATTTAGATTTATGATGACT 1784

RESULT 12
BD170980 1356 bp DNA linear PAT 17-JAN-2003

LOCUS BD170980
DEFINITION Process for producing prenyl alcohol.
ACCESSION BD170980.1 GI:27876792
VERSION WO 02053746-A/22.
KEYWORDS Saccharomyces cerevisiae (baker's yeast)
SOURCE Saccharomyces cerevisiae
ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 1356)
Oto, C., Obara, S., Muramatsu, M., Nishi, K. and Totsuka, K.
Process for producing prenyl alcohol
Patent: WO 02053746-A 22 11-JUL-2002;
TOYOKA MOTOR CORP. CHIKARA OTO, SHUSEI OBARA, MASAYOSHI MURAMATSU,
KIYOHICO NISHI, KAZUHIKO TOTSUKA
OS Saccharomyces cerevisiae (yeast)
PN WO 02053746-A/22
PD 11-JUL-2002

PF 20-DEC-2001 WO 2001JP011214
 PR 28-DEC-2000 JP 00P 403067
 PI CHIKARA OTO, SHUSEI OBATA, MASAYOSHI MURAMATSU, KIYOHICO NISHII,
 PI KAZUHIKO TOTSUKA
 PC C12N15/52, C12P7/04, C12N1/19, C12N1/21
 CC Process for producing prenyl alcohol
 FH Key
 FT source
 FT 1.1356
 Location/Qualifiers
 1.1356
 /organism="Saccharomyces cerevisiae (yeast)"
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"

ORIGIN

Query Match 10.2%; Score 132; DB 6; Length 1356;
 Best Local Similarity 50.5%; Pred. No. 4.7e-19;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

3 GTCAAAAGCATTTAGTGCACCTGGAAGAAAGCATTTCTGCTGGTGATATTGGTTCTTGA 62
 9 GTTGAGAGCTTCAAGGCCCGGAGGAAAGCGTTACTAGCTGGATATTAGTTTAA 68
 63 GCCAATTATGATGCTTATGACAGCATTTGATCAGATGCAATGATGCAAGTTAAACCC 122
 69 TACAAATATGAGCATTTGTAGTGGATTTATCGGCAAGAAATGATGCTGATGCCATCC 128
 123 AAAAGAAC-----CAGTTGAAAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTGC 176
 129 TTACGGTTCATTGCAAGGGCTGTGATTAAGTTGAAGCGGTGAAAAGTAAACAATTAA 188
 177 AAACGAGAAATGGAATATCATATCATCAATAACAG--AGAACCCGAGAAAGTTCA 233
 189 AGATGGGAGTGGCTGTACATATAGTCTTAAAGTGGCTTCTCTGTTCCATAG 248
 234 GTACGCGATTAATTCATTTTGAAGCACTATATTCATCGTTTGAAGTTTAAACCC 293
 249 CGGATCTAAGAACCTTTTCAATGAAAAGTATCGTAAAGTATTTAGCTTAAAC 308
 294 GACCGAAGATTTGATCTTGA---TCAATCTTATCTCAGACCTCGAATATCA 344
 309 TTAACATGACGACTCTGCAATGAAAAGTTGCTGTTATGATTTCTGATGATGC 368
 345 TTCACAGAAATGATCTGAAACCAAGACATCTCGAATGAGAAAAGAAATTTCTTACA 404
 369 CTACATTTCTCAGAGAGATAGCTTACGAAATCGTGCACAGAAAGTTGAATTTCA 428
 405 TTCTCGTGCATTAACGAAAGTGAAGAAAGCCGATTAAGTTCAATCGGACAGATTAGTTC 464
 429 TTCGACAGAAATGAAAGATTTCCCAAAACAGGGCTGCTGCGCAGGTTTATGAC 488
 465 AGTTGTGCGACAAAGTTTATATCCATTTTATCC-----CAATGTTATCAGTAC 515
 489 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATGCACTGGAAGAAATATGATACAA 548
 516 GAATTAAGATTTTGGACAAAGCTGCAAGATTTGCAATGTTTATGAGCCCAAAAAGAT 575
 549 ATATAGAGAAATTTATCATATTTTACCAAGTTGCTCATGTGCAAGCTCAGGTTAAAT 608
 576 AGATCTGGGTTGATGTTGCAACTGCAATTTATGCTGATTTGATATAGAAATTTGA 635
 609 TGGAGCGGGTTGATGATGAGCGGCGGACATATGATCTATCAGATATAGAAAGATTTCC 668
 636 GCCAGTTTGAATATGACGTGTTTCAAGTTCTAGAAAGTATCTGAGAAAGTTCCCCAC 695
 669 ACCCGATTAATCTCTAATTTGCGAGATTT-----GGAAGTCTACTTACCGCGATA 722
 696 AGAGTTGAAAAAATGATGAAAGTAACTGAGAAATCAAAATGAAAGATGATCAATTACC 755
 723 ACTGGGCGCTTTGGTTGATGAAAGACGTGAAATTTAGCAATTAAGTAACATTTACC 782
 756 ATACGAATCAAGTATTAATGAGTGAAGTCAAGGGTGGCTCAGAAACACCCAAATTTGT 815

DB 783 TTCGGATTAACCTTATGATGATGGCGCATATTAAGAAATGTTAGAAACGTAACACTGT 842
 816 ATCAGCAATCTCCATGGAAGAAAGGAGGAGAAAGAGCTGTTGTATGACCA 875
 DB 843 CAGAAAGTAAAGAAATTTGATGATTTGCGATATGCGCAAGAAAGCTTGAATATATACAA 902
 876 GCTTAATAGTCCAAATTTACGATTTATGAGGAAT 911
 DB 903 ACTGCATCATGCAAAATTTCTAGATTTATGATGATGACT 938

RESULT 13

BD171098 1356 bp DNA linear PAT 17-JAN-2003
 BD171098
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE

1 (bases 1 to 1356)
 Ota,C. and Obata,S.
 Process for producing prenyl alcohol
 Patent: WO 02053747-A 22 11-JUL-2002;
 TOYOTA MOTOR CORP, CHIKARA OTO, SHUSEI OBATA
 OS Saccharomyces cerevisiae (yeast)
 PN WO 02053747-A/22
 PD 11-JUL-2002
 PF 20-DEC-2001 WO 2001JP011215
 PR 28-DEC-2000 JP 00P 401701, 28-DEC-2000 JP 00P 403067 PR
 18-SEP-2001 JP 01P 282978

AUTHORS

CHIKARA OTO, SHUSEI OBATA
 PI CHIKARA OTO, SHUSEI OBATA
 PC C12N15/52, C12P7/04, C12N1/19, C12N1/21
 CC Process for producing prenyl alcohol
 FH Key
 FT source
 FT 1.1356
 Location/Qualifiers
 1.1356
 /organism="Saccharomyces cerevisiae (yeast)"
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"

COMMENT

1 (bases 1 to 1356)
 Ota,C. and Obata,S.
 Process for producing prenyl alcohol
 Patent: WO 02053747-A 22 11-JUL-2002;
 TOYOTA MOTOR CORP, CHIKARA OTO, SHUSEI OBATA
 OS Saccharomyces cerevisiae (yeast)
 PN WO 02053747-A/22
 PD 11-JUL-2002
 PF 20-DEC-2001 WO 2001JP011215
 PR 28-DEC-2000 JP 00P 401701, 28-DEC-2000 JP 00P 403067 PR
 18-SEP-2001 JP 01P 282978

FEATURES

1.1356
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"

ORIGIN

Query Match 10.2%; Score 132; DB 6; Length 1356;
 Best Local Similarity 50.5%; Pred. No. 4.7e-19;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

3 GTCAAAAGCATTTAGTGCACCTGGAAGAAAGCATTTCTGCTGGTGATATTGGTTCTTGA 62
 9 GTTGAGAGCTTCAAGGCCCGGAGGAAAGCGTTACTAGCTGGATATTAGTTTAA 68
 63 GCCAATTATGATGCTTATGACAGCATTTGATCAGATGCAATGATGCAAGTTAAACCC 122
 69 TACAAATATGAGCATTTGTAGTGGATTTATCGGCAAGAAATGATGCTGATGCCATCC 128
 123 AAAAGAAC-----CAGTTGAAAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTGC 176
 129 TTACGGTTCATTGCAAGGGCTGTGATTAAGTTGAAGCGGTGAAAAGTAAACAATTAA 188
 177 AAACGAGAAATGGAATATCATATCATCAATAACAG--AGAACCCGAGAAAGTTCA 233
 189 AGATGGGAGTGGCTGTACATATAGTCTTAAAGTGGCTTCTCTGTTCCATAG 248
 234 GTACGCGATTAATTCATTTTGAAGCACTATATTCATCGTTTGAAGTTTAAACCC 293
 249 CGGATCTAAGAACCTTTTCAATGAAAAGTATCGTAAAGTATTTAGCTTAAAC 308
 294 GACCGAAGATTTGATCTTGA---TCAATCTTATCTCAGACCTCGAATATCA 344

```

Db      309 TAACATGACGACCTACTGCAATAGAAAATTGTGTTATTGATATATTTCTGTATGATGTC 368
Qy      345 TTCAAGAAGATATCTGAACCAAGACATCTCGAATGAGAAAACATTTCTTTTACCA 404
Db      369 CTACCATTTCTGAGAGATGACGCTTACCGAACATGCGGCAACGAAAGATTGAGTTTCA 428
Qy      405 TTCTGTCGCACTTACCGAAGTGGAAAAGACCGATTAGTTTCATCGCAGGATTAAGTGC 464
Db      429 TTCCACAGAAATGAAAGAGTTCCCAAAACAGGGCTGGCTCTCGGCAAGTTTACGAC 488
Qy      465 AGTGTGGCCCAAGTTTATATATCCATTTTATCCC-----CAATGTATCAGTAC 515
Db      489 AGTTTAACTACACCTTGGCTCTCTTTTGTATCGACCTGGAAATATGATAGCAAA 548
Qy      516 GAATTAAGATATTTTGGACAACGTTGCAAGATTGCAATTTGTAAGCCCAAAAAAGAT 575
Db      549 ATATAGAGAAAGTTATTTATATTTAGCAAGTTGCTCATTTGTCAGAGCTGAGGTAAT 608
Qy      576 AGGATCTGGGTTGATTTGCCAATGCAATTTATGCTGATTGTATATAGAAATTTCA 635
Db      609 TGGAGGCGGTTGATGATGCGGCGGACATATGATATCATATATAGAAATTTCC 668
Qy      636 GCCAGCTTTGATTAATGACGTTTTCAGGTTCTAGAAAGATCTCGAAGATTCCCCAC 695
Db      669 ACCCGATTATCTCTTATTTGCCAGATATTT-----GGAAAGTCTACTTACCGCAGTAA 722
Qy      696 AGAGTTGAAAAAATGATTGAAAGTAACTGGGAATTCAAACATGAAGAATGTACATTACC 755
Db      723 ACTGGCGCATTTGTTGATGATGAAGAAGCTGGAATATTTAGCATTAAAGTACCATTTACC 782
Qy      756 ATGCGAATCAATTTTATATGCGTACGTCGCAAGGTTGCTCAGAAAACCCCAATTGCT 815
Db      783 TTGGGATTTACTTTTATGATGAGGCGCATTTTAAAGATGTTTCAGAAAACAGTAAACTGCT 842
Qy      816 ATCAGCAGTACTCCAAATGGAAGAAAAGCAAGAAAGAGCTGTTGTGATAGACA 875
Db      843 CCAGAAAGTAAAAAATTGTAATATTCGCAATATGCCAAGAAAGCTTGAATAATATACAGA 902
Qy      876 GCTTAATAGTGCATTTACAGTTTATGAAGAATT 911
Db      903 ACTGATCATGCAAAATTTAGATTATGATGATGACT 938

RESULT 14
SC9959 40397 bp DNA linear PLN 11-AUG-1997
LOCUS S.cerevisiae chromosome XIII cosmid 9959.
DEFINITION 249939.1 GI:887599
VERSION 249939.1 GI:887599
KEYWORDS dihydrofolate reductase; ERG8; MRE11; MRLP44; MTF1;
phosphomevalonate kinase; ribonuclease H; ribosomal protein L44;
RNH1; transfer RNA-Arg; ubiquitin carboxyl-terminal hydrolase.
SOURCE Saccharomyces cerevisiae (Baker's Yeast)
ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 16981 to 22170)
AUTHORS Venema,J. and Tollervey,D.
TITLE RRP5 is required for formation of both 18S and 5.8S rRNA in yeast
JOURNAL EMBO J. 15 (20), 5701-5714 (1996)
MEDLINE 97051828
PUBMED 8896463
REFERENCE 2 (bases 1 to 40397)
AUTHORS Skelton,J. and Churcher,C.M.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 40397)
AUTHORS Barrell,B., Rajandream,M.A. and Walsh,S.V.
JOURNAL Direct Submission
TITLE Submitted (22-JUN-1995) Saccharomyces cerevisiae chromosome XIII
sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge
CB10 1RQ E-mail: barrell@sanger.ac.uk
COMMENT Notes:
All CDS over 100 codons have been analysed. CDS that are completely

```

FEATURES

source

overlapped and those that are overlapped by more than 50% of their length by a larger CDS have been omitted from this analysis.
Details of the omitted CDS are available on request. The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons and the calculated codon adaptation index (CAI) is given for each CDS.
Cosmid 9959 overlapped at 5' by cosmid 8261, emb1 entry SC8261X, accessionno.249809 and at the 3' by cosmid 9408, emb1 entry SC9408, accession no.248756.

Location/Qualifiers

CDS

```

1. 40397
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="AB972"
/db_xref="taxon:4932"
/chromosome="XIII"
/map="13R"
/clone="cosmid 9959"
<1. 3156
/note="YM9959.01, len: > 1051, CAI: 0.15, incomplete
unknownorf, overlaps and extends cosmid 8261 orf
YM8261.13"
/codon_start=1
/product="unknown"
/protein_id="CA90190.1"
/db_xref="GI:887600"
/db_xref="GOA:Q03661"
/db_xref="Swiss-Prot:Q03661"
/translacion="SSSSEIIEVEKSEKIDGSTEKEIVPLSTDTTINNSLIGMESI
VSLSDADDAISENLTDPIMEIKTTPRYVIESVYSSTSYSDNTVAMPPOVEYVSP
FMNDPRNSLNDVYREKHDLLKSTLAALAPFKDAEIVAGVYKSLTSTSGHTNIF
HITSKETIQVSDLDSETEVTFEVENETGDKNOKSKPEGVANSTDESDTEKTFYS
AINTNTVGTDSCEDIJETASVNEMLRYCEKMNAMKMSGDECVKQNDGSKTQIS
FSTDSPNFOESNDNTEFSSTKYKVNSLDEDESKELTKAEVLDKDEESSESY
EODYADPEPNDGSENNIYKTKKDTLGIVEPKNVKNVHREETLFEANVSSVNV
ONKMDHTDVINOANVEYGERKYYIONTPTREAHSTIERIDENAINGMETIPRS
CVKTHNEVLFERRATTIENRTKALNNNTNMHDVSOACSDSDSDVSTABKVEGSAK
HNIDIRSSSEISIEVEPLKESDRSNIFSPPIRVIAVVGKGVADVABSPFKYKID
MDSSEDNVDVIGDYNODIFENKNSSTDAVNMKVSKEKSDDEDAVILGVTAAEAMN
DNGNSNRVINIDPTNGAYEEDSEVPOQVKENLKSSEPLVEGLQSEHREKDH
SENEBEEDTLYGDTISANISHNAPDDIKRQQLKNSIDENYSORLIEDSRKGNE
SDENVTSEKRDITFEKSVNEKYGALIEEDTFSRLDISIOHREEDLDSNNOREIE
EINSSEPRATVELLETGPTETLASKKNDDERORNTISTDIPSPPEKDEYDTSY
PVNSSEPRATVELLETGPTETLASKKNDDERORNTISTDIPSPPEKDEYDTSY
RSRHLSSHVDNEPHUNSTINIKVSESEPHQAVDIPVVEVKEDEGEMPEKSVLEQ
KPSMELINDSSPENNDSETNKDKTKAKKSKRKNVNSRRKKKITEGSSAASNT
KRRRGHEPKSRGQNTHPDYVK"
1. 312
/note="overlap with cosmid 8261, EMBL SC8261X, 249809,
positions 27890 to 28201"
3363. 4718
/gene="ERG8"
3363. 4718
/gene="ERG8"
/note="YM9959.02, ERG8 gene, len: 451, CAI: 0.15,
SW:ERG8_YEAST_P24521, phosphomevalonate kinase; contains
PS00627_GHMP kinases putative Afp-binding domain, conflict
at C-terminus due to a frameshift with ERG8_YEAST"
/codon_start=1
/product="Erg8p"
/protein_id="CA90191.1"
/db_xref="GI:887601"
/db_xref="SCD:50004833"
/db_xref="GOA:P24521"
/db_xref="SGD:50004833"
/db_xref="Swiss-Prot:P24521"
/translacion="MSELRASFAPGKALAGGYLVDTKYEAFFVGLSARMAVAHPY
GSLQSGDKFEVVKSKQFQDGBMLYHISPSGFIIVSGSKNPTEKVINAVSFYK
PMNDYCNRNLFVIDIFSDAYHSGEDSVTEHGRGRSLSPSHRTIEVPKTKGSSAG
LVTVLITLALSPFVSDLENNYDKREVYIHNLAQVAHCAQAGKIGSGFPVAAAAGSIR
YRRFPALISNPIDGSAITGSKLAHLVDEEDMNITISNHLPSGLTLMQDINKGSE
TVKLVOYKKNWYDSHPESLKIYTELHDHANSRPMDGSLKDLRLHRTHDYSDQIFESL

```

misc_feature

CDS

ERNDCQKYPEITVRDAVATIRRSFRKTKESGADIEPPVQTSLLDDCOTLKGLVT
CLIPGAGGYDAIAVITKOPVDIAQOTANDKRFSKYOMLWDTQADWGVREKDEPETYLD
K"

/gene="ERB8"
/note="P800627 GHMP kinases putative ATP-binding domain"
/complement(4978..6492)
/note="YW9599.03c, unknown, len: 504, CAI: 0.15, possible
membrane protein"
/codon_start=1
/product="unknown"
/protein_id="CAA90192.1"
/db_xref="GI:887602"
/db_xref="GOA:P32829"
/db_xref="SGD:S0004837"
/db_xref="SwIsa8-Prot:P32829"

CDS

/translation="MTSTRTTLYAQAACACIACIIPSAIIPPAKAPLIEGYTHE
LDPDGRDLCTQADLKINFALSAVATNTIMALPVGKILDMYGRVCGIIGSCILF
LASGNFISAKHLVSLMDPYLVGTITLAVAGPVPFISCPQLANSFQRCSTVALTLGTS
PSSSALFLYLYLYQNPPTLVNVRFPFLYLVIPVFLACQLTIMPSSKYTNHIA
KIAVGLDNGRLIGDGTSGGIIIPDEQROSLAIEREDSIISRPOKRSVLETYE
DKLQKSGGIRGVHLGKSAVROIKSPWFLMLPALVAMLRINYPATRTQREYILN
DPDLAKLNSIFDMLPLGGAVSIFPIGLLHDITLTLITLFTSTATGVGLIPN
SFTWNLIGVLLVVRPFTYTVSDYSKVFEPDFTGVYGLSCIGIFNNSQMLD
KMTHTFMNPPINLIVILTVFVSLTLTPYRSQILPKPNERGLSSNYQTI"
/complement(6685..7356)
/note="YW9599.04c, unknown, len: 223, CAI: 0.10, similar
to SW:DYR_SCHPO_P36591, dihydrofolate reductase [29.4%
identity in 218 aa overlap]"
/codon_start=1
/product="unknown"
/protein_id="CAA90193.1"
/db_xref="GI:887603"
/db_xref="SGD:S0004835"
/db_xref="SwIsa8-Prot:P05015"

CDS

/translation="MTKYNLMLGLAQSGDYASKTKGRFAEMKLGKLYYPYAPNE
PPADVPDLGVIADAPGDGENTGVLALENDPSTGGYFIPQTTIDYLVNLRGEP
FAGIVGFSQAGVAGIATFNGLLGLTEEQPPLFPAVAGSPQOYOEOYDLH
PIVSPLHYQSELDTITTEPAKQGLYNSTEDSRILMHSGHFPVNSGFPKRYAQM
LQQLT"
/note="YW9599.05, unknown, len: 471, CAI: 0.17, similar to
putative ubiquitin carboxyl-terminal hydrolases, eg
SW:UBPX_HUMAN_P40818, probable ubiquitin
carboxyl-terminal hydrolase [31.5% identity in 352 aa
overlap]; contains P800973 Ubiquitin carboxyl-terminal
hydrolases family 2 signature 2"
/codon_start=1
/product="unknown"
/protein_id="CAA90194.1"
/db_xref="GI:887604"
/db_xref="GOA:P50102"
/db_xref="SGD:S0004836"
/db_xref="SwIsa8-Prot:P50102"

misc_feature

gene

/translation="MSICPHIOOVFONKESKDGVLTKTCAARYILNHSVPEKFLNTM
KCGTGEINSGATFMCLOCGFGCMNHSPLSHSKOIGIPIGNSNGLLFPFKEDY
IGNLIDILNAILAKYMDYDCTKTPVPSMRKRGSLGIMSGTCRPSLLOCLINRPY
FIRHMSQTHSNCKVRSPDKFSCHLDKIVHELGAALTKQASSSTSTNKQCTIVY
LITCAMKINQNLAGYSQODAHFWQFIINQIHQSYVLDLPNKEVSRANKQCECIVH
TVFEGSLSESLVPCGONNSKTTIDPFLDLSLIDKDKLTYELDFHKEKQLOCEFNH
HCGENSTQDAIKQGIHKLPSVLVQLRPFHLNGLSRRKLDLPLEPTVIMKNKY
STKEKDSKSENGKVPDIYELIGIVSHKTVNEGHYIACCKISGGWMPFNDSWVSSI
SOEVLKROAYLFTYIRQVN"
/note="P800973 Ubiquitin carboxyl-terminal hydrolases
family 2 signature 2"
/complement(9622..11700)
/gene="MRB11"
/complement(9622..11700)
/gene="MRB11"

/note="YW9599.06c, MRB11 gene, len: 692, CAI: 0.14,
SW:MR11_YEAST_P32829, required for repair and meiotic
recombination, conflict, this sequence is longer due to a

Frameshift with SW:MR11_YEAST"
/codon_start=1
/product="Mrb11p"
/protein_id="CAA90195.1"
/db_xref="GI:887605"
/db_xref="GOA:P32829"
/db_xref="SGD:S0004837"
/db_xref="SwIsa8-Prot:P32829"
/translation="MDYPPDPIRILITITLDNHNAGYENDPIRGDSSWKTFFHEVMILAK
NNNVAVOSGDLPHVNRKPSKSLVQVTKTLRIGCMGKPCLELIDSPGVNHYDEF
TNVYEDENPNISIPVPSISGHNDASDLSLCPMDILAHNLINHPKGVIESDKIKY
VPLPQKSTKLALGLAVRDERLFRFKDGVTFEYPTNKGSEFNLMCTQONHTG
HTNTAFLEQPLDPLDMVIMGHEECIPNIVHNPDKPDVQPSGSAVSLCEAEQ
PKYVFLIDIKYGEAPKATPIPLETIRTKMSISIQDYPHLRPHKDAISKYLIEQVE
EMIRANETKOKLADQSGMVNLEPRLIRLVYDAPASPTQSPIDYOVENPRFS
NRPFVRVANGNVVQFYKRSKSPVRSKNGSTISDRDKEFLPSBSGSELEVOYTLV
NDLNRKQSLPVEGLNBAVYKFPDQKREKRLAKEPFISHISNBVGILSTNEPRTD
DAEEMKALIKOYKRANSVPRPEKNDSTNFAFNQNGDSFRSREVRFGSPDITQDS
HYDNESRITTHSQAESRPSKPKRKRTAKKRIPAFSDSTVISAENELGNNDAQD
DVDLIDENRIIMVSTDEEDASVGLNLNGRKTITRPAASTKTASRRKGRASRTPTDI
LGLSLAKRKK"

/complement(join(12007..12245,12393..12450))
/gene="MRPL44"
/complement(join(12007..12245,12393..12450))
/gene="MRPL44"
/note="YW9599.07c, MRPL44 gene, (YMR44), len: 98, CAI:

Query Match		10.2%; Score 132; DB 8; Length 40397;
Best Local Similarity		50.5%; Pred. No. 3,4e-19;
Matches 473;		Conservative 0; Mismatches 430; Indels 33; Gaps 5;
QY	3	GTCAAAAGCATTTAGTGCACCTGAAAGCATTTCTGCTGGTGTGATATTTGTTCTTGA 62
DB	3371	GTTCGAGGCTTCAAGTCCCGGAGAAAGCATTTAGTGTGATATTTGTTCTTGA 3430
QY	63	GCCAAATTAATGATGCTTATGACAGCATTTGCATCAGCATGATGATGATTAACAC 122
DB	3431	TACAATAATGAAAGCATTTGATGCGATTAATGCGCAAGAAATGATGCTGATGCCATCC 3490
QY	123	AAAAGAAC-----CAGTTGAAAAGATCTGAATCAAAATTTCTTCAACCCCAATTGCG 176
DB	3491	TTACGCTTCATTCGAAAGGCTGATPAAGTTTAAAGTCGTGAAAAGTAAACAATTGA 3550
QY	177	AAACGAGAATGGGAATATCATATTCATCAATATACAG---AGAAAGCCAGAAAGTTCA 233
DB	3551	AAGTGGGAGTGGCTGTACATATTAAGTCTTAAAGTGGCTTCATTCCTGTTCCATAGG 3610
QY	234	GTACAGCATTAATTCATTTTGAAGCACTATATTCATGTTTGAAGTTATATTCAC 293
DB	3611	CGATCTAAGAACCTTCATTTGAAGAAAGTTATCGTAAAGTATTTAGTACTTAAACC 3670
QY	294	GACCGAAGCATTTGATCTTGAAG-----TCATCATTTTCTAGACCCCTGATATCA 344
DB	3671	TAAACATGACGACCTACTGCAATGAAGAACTGTTCGTATTTGATTAATTTCTGATGATGC 3730
QY	345	TTCAACAAGAAATACGAAACCAAGCAATCTCGAATGAGAAAGAAACATTTCTTACCA 404
DB	3731	CTACATTTCTCGAGAGATAGCTTACCGAACAATGTGCAACAGAAAGTTAGTTTCA 3790
QY	405	TTCTGTGTCATTAACGAAAGTGAAGAACCCGATTAAGTTGATCGGACAGATTAAGTGC 464
DB	3791	TTTCGACAGAAATTTGAAGAAAGTTCCCAAAACAGGGCTGGGCTCTCGGACAGTTAGTCA 3850
QY	465	AGTTGTGTCACAAAGTTTATATCCCATTTATCC-----CAATGTTATCAATGC 515
DB	3851	AGTTTAACTACAGCTTGGCCCTCTTTTGTATCGACCTCGAAGAAATTAATGTAGACA 3910
QY	516	GAATTAAGATATTTGACAAAGTTGACAGATGACATTTGTTAGCCCAAAAAAGAT 575
DB	3911	ATATGAGAAATTTATCATATTTAGACAAAGTTGCTCATTTGCAAGCTCAGAGTAAAT 3970
QY	576	AGGATCTGGGTTGATGTTGCAACTGCAATTTATGCTGATTTGATATAGAAATTTCA 635

```

Db      3971 TGGAGCGGTTTATGATGACGGCGGACAGATATGATCTATCAGATATAGAAATGCC 4030
Qy      636 GCCAGCTTTGATTAATGACGTGTTTCAAGTCTTAGAAAGTATCCTGAGAGTTCCAC 695
Db      4031 ACCCGCATTAATCTCTAATTTCCAGATTT-----GGAAGTCTTACGCGACGTAA 4084
Qy      696 AAGATTGAAAAAATGATTGAAAGTAAGTAAGTCAACATGAAAGATGTATAC 755
Db      4085 ACTGGCGCTTTGGTTGATGAGAAAGACTGGAATTTTCAGATTAAGTAACATTTAAC 4144
Qy      756 ATACGATCAAGTATTATTAATGGTGAAGTCAAGGAGGCTCAGAAACCCCAATTGT 815
Db      4145 TTGGGATTAATCTTATGATGAGGAGATTTAGATGTTTCAAGAAACAGTAAATCTGT 4204
Qy      816 ATCAGAGTATCTCCATGAGAAAAAGAAAGCCAGAAAGAGCTGTGTGTATGACCA 875
Db      4205 CCAGAGGTTAAAAAATTTGATGATTCGATATGCCAGAAAGCTTGAATAATATACAGA 4264
Qy      876 GCTTAATATGTCGCAATTACAGTTATGAAAGAAATT 911
Db      4265 ACTGATCATCAAAATTTAGATTTATGATGACT 4300

```

RESULT 15
CR382131_07/c
WPCOMMENT

Sequence split into 43 fragments LOCUS CR382131 Accession CR382131

Fragment Name	Begin	End
CR382131_00	1	110000
CR382131_01	100001	210000
CR382131_02	200001	310000
CR382131_03	300001	410000
CR382131_04	400001	510000
CR382131_05	500001	610000
CR382131_06	600001	710000
CR382131_07	700001	810000
CR382131_08	800001	910000
CR382131_09	900001	1010000
CR382131_10	1000001	1110000
CR382131_11	1100001	1210000
CR382131_12	1200001	1310000
CR382131_13	1300001	1410000
CR382131_14	1400001	1510000
CR382131_15	1500001	1610000
CR382131_16	1600001	1710000
CR382131_17	1700001	1810000
CR382131_18	1800001	1910000
CR382131_19	1900001	2010000
CR382131_20	2000001	2110000
CR382131_21	2100001	2210000
CR382131_22	2200001	2310000
CR382131_23	2300001	2410000
CR382131_24	2400001	2510000
CR382131_25	2500001	2610000
CR382131_26	2600001	2710000
CR382131_27	2700001	2810000
CR382131_28	2800001	2910000
CR382131_29	2900001	3010000
CR382131_30	3000001	3110000
CR382131_31	3100001	3210000
CR382131_32	3200001	3310000
CR382131_33	3300001	3410000
CR382131_34	3400001	3510000
CR382131_35	3500001	3610000
CR382131_36	3600001	3710000
CR382131_37	3700001	3810000
CR382131_38	3800001	3910000
CR382131_39	3900001	4010000
CR382131_40	4000001	4110000
CR382131_41	4100001	4210000
CR382131_42	4200001	4224103

Continuation (8 of 43) of CR382131 from base 700001 (CR382131 Yarrowia lipolytica chromo

Query Match 7.8%; Score 101.6; DB 8; Length 110000;
Best Local Similarity 49.7%; Pred. No. 2e-12;
Matches 319; Conservative 0; Mismatches 314; Indels 9; Gaps 2;

```

Qy      394 TTCTTACCATTCCTGTCGCCATTACCGAAGTGAAGAAAGACCGATTGTTTCATGGCA 453
Db      1666 TTTCATACCAAAAAGGCGATCCACAGAGGCTCTTAAAGCGGGCTGTAGTCCGCT 16509
Qy      454 GGATTAAGTGTCAAGTTGTCACAAAGTTATATCCATTATTCATCCCATGTTATCAGT 513
Db      16608 GCTTTACCAACCGTCTTGTGGCAGCTTTGCT---CAAGTATACGGATGATTCCTTG 16552
Qy      514 ACGAATTAAGATTTTTCACAAGCTTGCACAGATTCACATTTGATGCCAAAAAAG 573
Db      16551 CATTAACCCACCTGTTCACAACCTGTCCAGGTGTCACCTCTCGGCACAGAAAG 16492
Qy      574 ATAGATCTGGGTTTGTATGTGTCACATTCATTTATGTCGATTTATATGAAAGTTT 633
Db      16491 ATTGGGTCTGGATTGAGTGGCTTCGGCCGTTTGTGCTCTTATGATGAGTTTC 16432
Qy      634 CAGCCAGCTTTGATTAATGACGTTGTTCAAGTTCTAGAAAAGTATCCTGAAAGTCCC 693
Db      16431 CCGCGGAGTCCGTAACATGTCATTGACAGTGAAGGA-----CCTCCGATACGGG 16378
Qy      694 ACAGATTGAAAAAATTTGATTAAGTAACTGCGGAATTCAAACATGAAGATGTACTTA 753
Db      16377 GCTGTGTTGAAACCTACCGTTATCAAAAGTGAAGTGACTGTGAACCATCTTCTTG 16318
Qy      754 CCAATACGGAATCAAGTTATTAATGAGTGAAGTCAAGGAGTGTGACAAACCAATTTG 813
Db      16317 CCGCGGGAATCAACCTGCTTAATGGAGAGTCAAGGAGATCTGAGATCCAGGTATG 16258
Qy      814 GTATCAGAGTACTCCATGAGAAAAAGAAAGCCAGAAAGAGCTGTGTGTATGAC 873
Db      16257 GTGGCCAGGTGATGATGATGCGAAGCAAGCAAGCCCGAGAAACCGAGATGTGTGAGA 16198
Qy      874 CAGTTAATATGTCGCAATTTCAGTTATGAAAGAAATGAGGAATGCGTGAATAAC 933
Db      16197 GATCTCAACGCTGCCAATCATGTCATGTCATGTTTCAAGCACTGCGCAAGCTCTT 16138
Qy      934 GACTCAGAACCCAGACTTATTAATAAGATTGATGATCACTTGTGAGCTTTGACTGTT 993
Db      16137 CTACATCAACAGAGGCTTACGAACAATTTGGCCGAGAGCTGCTCTTCAACGCTTA 16078
Qy      994 GCGATTAAAGATCAAGAAAAAGGTTACAAAGCTTAACAA 1035
Db      16077 AAGATGATATGTTGAGAAACCTCTCGAGAACTAGCAGCA 16036

```

RESULT 16
SPAC343/c 42947 bp DNA linear PLN 20-JUN-2003
LOCUS
DEFINITION S.pombe chromosome I cosmid c343.
ACCESSION AL109739
VERSION AL109739.1 GI:5706501
KEYWORDS
SOURCE Schizosaccharomyces pombe (fission yeast)
ORGANISM Schizosaccharomyces pombe
Bukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomyces
1 (bases 1 to 42947)

REFERENCE
AUTHORS
Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R.,
Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., Basham, D.,
Bowman, S., Brookes, K., Brown, D., Brown, S., Chillingworth, T.,
Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, P.,
Feltwell, T., Fraser, A., Gentles, S., Goble, A., Hamlin, N., Harris, D.,
Hidvegi, E., Hodgson, G., Holroyd, S., Hornsby, T., Howarth, S.,
Huckle, E., Hunt, S., Jags, K., James, K., Jones, L., Jones, M.,
Leather, S., McDonald, S., McLean, J., Mooney, P., Moulé, S.,
Munhall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neill, S.,
Pearson, D., Quail, M.A., Rabinowitch, E., Rutherford, K., Rutter, S.,
Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M.,

TITLE	JOURNAL	Squares, R., Squares, S., Stevens, K., Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckart, G., Aert, R., Robben, J., Grymoprez, B., Weltjens, I., Vanstrele, E., Rieger, M., Schaefer, M., Miller-Auer, S., Gabel, C., Ruch, M., Dusterhoft, A., Fritze, C., Holzer, E., Moestl, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambut, R., Purnelle, B., Goffeau, A., Cadieu, B., Dreno, S., Gloux, S., Lelaure, V., Mottier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Gallardin, C., Tallada, V.A., Garzon, A., Rhode, G., Daga, R.R., Cruzado, L., Umenae, J., Sanchez, M., del Rey, F., Benito, J., Dominguez, A., Revuelta, D.L., Moreno, S., Armstrong, D., Forburg, S.L., Cerutti, L., Lowe, T., McCombie, M.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Uesary, D., Barrell, B.G. and Nurse, P.
	PUBMED	The genome sequence of Schizosaccharomyces pombe
REFERENCE	2	Nature 415 (6874), 871-880 (2002)
AUTHORS	11859360	(bases 1 to 42947)
JOURNAL	Murphy, L., Harris, D., Wood, V., Rajandream, M.A. and Barrell, B.G.	Submitted (03-AUG-1997) European Schizosaccharomyces genome sequencing project, Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: pombe@sanger.ac.uk
COMMENT	Notes:	Details of S. pombe sequencing at the Sanger Institute are available on the World Wide Web. (URL, http://www.genedb.org/genedb/pombe/index.jsp) (URL, http://www.sanger.ac.uk/projects/S_pombe/) CDS are numbered using the following system eg SPAC343.01c. SP (S. pombe), A (chromosome 1), CSH10 (cosmid name), .01 (first CDS), c (complementary strand). However, clones may have been reorientated since the original submission, therefore the complementary strand notation may be invalid for strand inference. IMPORANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once.
FEATURES	Location/Qualifiers	1..42947
source	/organism="Schizosaccharomyces pombe"	
	/mol_type="genomic DNA"	
	/strain="972h"	
	/db_xref="taxon:4896"	
	/chromosome="I"	
	/map="IL"	
gene	/clone="cosmid C343"	
	/complement(join(61..1333,1400..1407))	
	/gene="SPAC343.01c"	
	/complement(join(61..1333,1400..1407))	
CDS	/gene="SPAC343.01c"	
	/note="phosphomevalonate kinase (predicted); similar to S. cerevisiae ERG8"	
	/codon_start=1	
	/protein_id="CAB52264.1"	
	/db_xref="GI:5706502"	
	/db_xref="GOA:Q9UT88"	
	/db_xref="TrEMBL:Q9UT88"	
	/translation="MKVTCGAPGKULIAGVIVLDPQSGVIGITAKGYASTLLTDKCGIVKVSPOFINAEWLNIDWTVSPIRHVIYNCELEKPNPFVQLAFYVNIYFESTGRPLCWDLOVTLQVNDNAYHQPOLKPDQTSYKPNFLACTLGGVHAKGLSSAAMTSLIGSLFLSLRLTLDIDGDKSLIDSTKVIYHNLAIQAHSAQGSVGDVGAATAGSCYLRDPDLIBOLIVYDEQIKINISFELKIVSKMSDVPPQLPATYCLIMDVAGSSRPGKVKVQOKQENPESKCPDILSRVLSINCPKLSLSLSELQSPRSIRIIRITIVKAVNDIEPKQRTIILNIDLEQIVGIVGAGGPAQPCLAINTHEIENVITKWDGVNPDVSPAFGLAVE"	
misc_feature	/complement(1334..1353)	
	/gene="SPAC343.01c"	
	/note="cctaactagccttgagcag, splice branch and acceptor"	
	/complement(1394..1399)	
misc_feature	/gene="SPAC343.01c"	
	/note="gtaagc, splice donor sequence"	
gene	2277..2825	
	/gene="SPAC343.02"	
	/note="mitochondrial ribosomal protein S19; similar to S. cerevisiae IMG1"	
	/codon_start=1	
	/protein_id="CAB52265.1"	
	/db_xref="GI:5706503"	
	/db_xref="GOA:Q9UT87"	
	/db_xref="TrEMBL:Q9UT87"	
	/translation="MFNAKHFNPGLGPGWLQKGIISLKNRPYKPPGPKHVDLSLEKCKSVYDEGSEKFMFRHSPNRYRGAVILVSYKSPKSDVNRAGVILRIHREPSSILRNWGVGVYLLPIYSPQIKRIVLKENGLSRPRAYLSTYLRPFRRLPPVSLVRKRIEQLQHKP"	
misc_feature	2484..2750	
	/gene="SPAC343.02"	
	/note="match to PF01245 Ribosomal_L19, Ribosomal protein L19 Score 24.79"	
	/join(3630..3648,3692..3814,4032..4174)	
gene	/gene="SPAC343.03"	
	/note="synonym: APC11"	
	/join(3630..3648,3692..3814,4032..4174)	
CDS	/gene="SPAC343.03"	
	/note="anaphase-promoting complex (APC) (PMID 12477395); involved in cyclin degradation (required); involved in metaphase-anaphase transition (required); zinc finger protein; zfc-C3HC4 type (RING finger); ubiquitin ligase (E3) (predicted); conserved eukaryotic protein; similar to S. cerevisiae APC11"	
	/codon_start=1	
	/protein_id="CAB52266.1"	
	/db_xref="GI:5706504"	
	/db_xref="TrEMBL:Q9UT86"	
	/translation="MKVILRYHAIANTWPTPKDQVIGICRVPEPGCCPOCTSPGDNCPYWGKKHIFHAICIQNLATSGSQGCGMDQTFVVAISTNKSRTQ"	
misc_feature	3649..3654	
	/gene="SPAC343.03"	
	/note="gtaagc, splice donor sequence"	
	/complement(3681..3691)	
misc_feature	/gene="SPAC343.03"	
	/note="cctaactagcag, splice branch and acceptor"	
	/complement(3815..3820)	
misc_feature	/gene="SPAC343.03"	
	/note="gtaaga, splice donor sequence"	
	/complement(4019..4031)	
misc_feature	/gene="SPAC343.03"	
	/note="cctaactagcag, splice branch and acceptor"	
	/complement(join(4455..5753,5878..6028,6083..6156))	
gene	/gene="SPAC343.04c"	
	/complement(join(4455..5753,5878..6028,6083..6156))	
CDS	/gene="SPAC343.04c"	
	/note="WD repeat protein"	
	/codon_start=1	
	/protein_id="CAB52267.1"	
	/db_xref="GI:5706505"	
	/db_xref="TrEMBL:Q9UT85"	
	/translation="VALDEKFOLEVIHLLQFLNDYGYDESILKALEKETGLVSETEVRIKQAVIQGDMITAEAFSIMOLRDSKREKAEOLKQKCLDLASBGAICAEIYVLQNFESTDPEKKEKRIIVSITLSNNKSNNEILTKNGYCTRDLNQLSEIYSPILLPKRLEHLIQQAKDVQSSQVHVHAKNSFLSDYKADSELPKSYHVFHSDSWPISYHNGVYILASASDKTALIFDVVNLKRVRLIGHIDTVAYIRMSPDRLVSCSDKSVILMDLPTKEDKLDYGGFSVSCCMPLGLSFTTSPDCHITRWMLNGSLIKWEDVNIYDMALTEKDGTLVYVGEPLINADKHAIGNFTRECIKISLOSQVTSICLSKDSKYALNLNLPHTTFLMDLEENLIRVQMGHKGIFLIGSGFGKDDTPVLSGEDDKIRIWRSGKLATLSHYVCNVAIVNPDPYQFASAGDNTYRISNNDNPRQ"	
misc_feature	/complement(4482..4601)	
	/gene="SPAC343.04c"	
	/note="match to PF00400 WD40, WD domain, G-beta repeat"	
	/complement(5130..5228)	
misc_feature	/gene="SPAC343.04c"	
	/note="match to PF00400 WD40, WD domain, G-beta repeat"	
	/complement(5256..5360)	
misc_feature	/gene="SPAC343.04c"	
	/note="match to PF00400 WD40, WD domain, G-beta repeat"	

```

misc_feature      complement (5754..5769)
                  /gene="SPAC343.04c"
                  /note="ctacatgctactctag, splice branch and acceptor"
misc_feature      complement (5872..5877)
                  /gene="SPAC343.04c"
                  /note="gtatac, splice donor sequence"
misc_feature      complement (6029..6046)
                  /gene="SPAC343.04c"
                  /note="ctacagctatcttgtag, splice branch and acceptor"
misc_feature      complement (6077..6082)
                  /gene="SPAC343.04c"
                  /note="gtatac, splice donor sequence"
gene              complement (6814..6884)
                  /gene="SPATRAGLY.01"
                  complement (6814..6884)
                  /gene="SPATRAGLY.01"
                  /product="tRNA-Gly"
                  /note="tRNA Gly anticodon GCC, Cove score 72.94"
                  /evidence=not experimental
                  join(7781..7783,7839..7938,7981..9737)
                  /gene="vma1"
                  /note="synonym: SPAC343.05"
                  join(7781..7783,7839..7938,7981..9737)
                  /gene="vma1"
                  /EC_number="3.6.3.14"
                  /note="ny-type ATPase; H(+)-ATPase (PMID 1441756); vacuolar
                  ATP synthase (catalytic subunit A) (PMID 1441756); similar
                  to S. cerevisiae Tfp1"
                  /codon_start=1
                  /protein_id="CAB52268.1"
                  /db_xref="GI:5706506"

Query Match      6.7%; Score 86.6; DB 8; Length 42947;
Best Local Similarity 56.5%; Pred. No. 5.1e-09;
Matches 205; Conservative 0; Mismatches 149; Indels 9; Gaps 2;

QY 499 CCCATGTTATCAGTACGAATTAAGATATTTTGCACACGTTGCACAGATTGCACATTTGT 558
DB 786 CTCAAAATTGATGATTAACCAAGTCATAGTTCACATTTGCACAAAATGCCCCATGT 727
QY 559 TATGCCCAAAAAAGATAGATCTGGTTGATTTGCAACTGCATTTATGTCGTATT 618
DB 726 TCTGCACAGAAAAGTAGAAGTGGTGTGATGTTGGCGCTGCCACTTGGGAAAGTTGT 667
QY 619 GTATATGAAGATTTCAGCCAGCTTTGATTAATGACGTGT-----TCAGTTCTAGAA 672
DB 666 ATTTATGAAGTGTGATCTTAATTAATTGAACAACCTGCTGCCATPACGACGACAA 607
QY 673 AGTATCTCGAAGAGTTCCCAACAGAGTTGAAAAAATTGATGAAGTAAGTAAGTAATTC 732
DB 606 ATTAAGAACATTAATTTTTCACCGAAGTGAAGAAAAATAGATTTCTAAAAAATGG---TCA 550
QY 733 AAACATGAAGAATGTACATTACCATACGGAATCAAGTATTAAATGAGGAGTCAAGGT 792
DB 549 GATGTTGTTCCATTCAATTTGCTGCTACTTATTTGCTTAATGGGGAGTGTGCTGT 490
QY 793 GGCTCAGAAACCCAAATTTGATCAAGAGTATCCAAATGAAAAAGAAAAAGCCAGAA 852
DB 489 GGCAGTAGACACCGGGTATGTGTAATAAAAGTAGACGACATGGCAAAAAGAAATCCAGAG 430
QY 853 GAA 855
DB 429 GAA 427

```

```

RESULT 17
AX485723/c
LOCUS
DEFINITION Sequence 3023 from Patent WO02053728.
ACCESSION AX485723
VERSION AX485723.1 GI:22319939
KEYWORDS

```

```

SOURCE
ORGANISM Candida albicans
REFERENCE
AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Olsen,K.L.
TITLE Gene disruption methodologies for drug target discovery
JOURNAL Patent: WO 02053728-A 3023 11-JUL-2002;
          Bilitra Pharmaceuticals, Inc. (US)
FEATURES
source
  1..65
    /organism="Candida albicans"
    /mol_type="unassigned DNA"
    /db_xref="taxon:5476"
ORIGIN

```

```

Query Match      5.0%; Score 65; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.00663;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTTAGTGACCTGGAAGCAATTTCTGCTGCTGATATTGTTCTT 60
DB 65 ATGTCAAAAGCATTTAGTGACCTGGAAGCAATTTCTGCTGCTGATATTGTTCTT 6
QY 61 GAGCC 65
DB 5 GAGCC 1

```

```

RESULT 18
AX241110/c
LOCUS
DEFINITION Sequence 348 from Patent WO0160975.
ACCESSION AX241110
VERSION AX241110.1 GI:15797985
KEYWORDS
SOURCE
  1. synthetic construct
  2. synthetic construct
  3. artificial sequences.
REFERENCE
  1. Roemer,T., Jiang,B., Boone,C. and Bussey,H.
  Gene disruption methodologies for drug target discovery
  Patent: WO 0160975-A 348 23-AUG-2001;
  Bilitra Pharmaceuticals, Inc. (US)
FEATURES
source
  1..90
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /note="DNA primer"
ORIGIN

```

```

Query Match      5.0%; Score 65; DB 6; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.00661;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTTAGTGACCTGGAAGCAATTTCTGCTGCTGATATTGTTCTT 60
DB 65 ATGTCAAAAGCATTTAGTGACCTGGAAGCAATTTCTGCTGCTGATATTGTTCTT 6
QY 61 GAGCC 65
DB 5 GAGCC 1

```

```

RESULT 19
CNS06KLC
LOCUS
DEFINITION T3 end of clone AT0AA006B09 of library AT0AA from strain CBS 4311
of Saccharomyces servazzii, sequence tagged site.
ACCESSION AL403030
VERSION AL403030.1 GI:12162807
KEYWORDS

```

SOURCE Saccharomyces servazzii
ORGANISM Saccharomyces servazzii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 1 to 963)
AUTHORS Soucier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bollotin-Pukhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Maupertuy,A., Neuvéglise,C., Olier-Kalogeropoulos,O., Potier,S.,
Sautin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Winker,P. and Weissbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
PUBMED 11152876
REFERENCE 2 (bases 1 to 963)
AUTHORS Casaregola,S., Lepingle,A., Bon,E., Neuvéglise,C., Nguyen,H.,
Artiguenave,F., Winker,P. and Galliardin,C.
TITLE Genomic exploration of the hemiascomycetous yeasts: 7.
Saccharomyces servazzii
JOURNAL FEBS Lett. 487 (1), 47-51 (2000)
MEDLINE 20584717
PUBMED 11152882
REFERENCE 3 (bases 1 to 963)
AUTHORS Genoscope.
TITLE Direct Submision
JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
bege@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This STS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
source location/Qualifiers
1..963
/organism="Saccharomyces servazzii"
/mol_type="genomic DNA"
/strain="CBS 4311"
/db_xref="taxon:27293"
/clone="AT0AA006B09"
/clone_1fb="AT0AA"
/note="end : T3"
misc_feature <817..>954
/note="similar to Saccharomyces cerevisiae ORF YMR220w [
ERG8 : phosphomevalonate kinase]"
/evidence=not_experimental
ORIGIN
Query Match 4.3%; Score 56; DB 11; Length 963;
Best Local Similarity 66.7%; Pred. No. 0.051;
Matches 80; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 4 TCAAAAGCATTTAGTGCACCTTGAAAGCATTTCTTGCTGATATTGTTCTTGAG 63
DB 832 TCGAGAGCATTTAGTACCCCGGAAAGCCCTTATTATGCTGGATCTGTTTGAT 891
QY 64 CCAATTATGATGCTTATGTCACAGCATTTGATACGAAATGATGATTTATACACA 123
DB 892 CCACATTAACAGTCATTCTGTAGCTCTTTCAGCGAAGATGATGCTGTGTAAGACA 951

ACCESSION AL401081
VERSION AL401081.1 GI:12158232
KEYWORDS STS.
SOURCE Saccharomyces bayanus
ORGANISM Saccharomyces bayanus
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 1 to 1037)
AUTHORS Soucier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bollotin-Pukhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Maupertuy,A., Neuvéglise,C., Olier-Kalogeropoulos,O., Potier,S.,
Sautin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Winker,P. and Weissbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
PUBMED 11152876
REFERENCE 2 (bases 1 to 1037)
AUTHORS Bon,E., Neuvéglise,C., Casaregola,S., Artiguenave,F., Winker,P.,
Aigle,M. and Durrens,P.
TITLE Genomic exploration of the hemiascomycetous yeasts: 5.
Saccharomyces bayanus var. uvarum
JOURNAL FEBS Lett. 487 (1), 37-41 (2000)
MEDLINE 20584715
PUBMED 11152880
REFERENCE 3 (bases 1 to 1037)
AUTHORS Genoscope.
TITLE Direct Submision
JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
bege@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This STS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
source location/Qualifiers
1..1037
/organism="Saccharomyces bayanus"
/mol_type="genomic DNA"
/strain="CLIB 533"
/variety="uvarum"
/db_xref="taxon:4931"
/clone="AS0AA025E12"
/clone_1fb="AS0AA"
/note="end : T7"
misc_feature <3..>560
/note="similar to Saccharomyces cerevisiae ORF YMR220w [
ERG8 : phosphomevalonate kinase]"
/evidence=not_experimental
misc_feature complement(<786..>1005)
/note="similar to Saccharomyces cerevisiae ORF YMR221c [
weak similarity to photolysate II protein D2]
1 putative frameshift(s)"
/evidence=not_experimental
ORIGIN
Query Match 4.3%; Score 55.6; DB 11; Length 1037;
Best Local Similarity 57.5%; Pred. No. 0.063;
Matches 100; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 986 TGACCTGGCATTTAGAACATCAGAAAGGCTTACAGCATTAACAAATCAGAGC 1045
DB 256 TGAAGATTTCTGTGACAGATTAAGCCTGTTTACAGAAATTAACATTAAGATCTGAG 315
QY 1046 TTCCAATTGAACCTGATGTCAAACCCAGTTGAGACCGTTGCAAGATTCCTGTT 1105

AUTHORS	Sawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Anaki, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, T., Nishijima, K., Otsuka, R., Nakazawa, H., Takamiya, M., Kato, Y., Yoshizawa, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kusihida, N., Oguchi, A., Aoki, K., Masuda, S., Yanagii, M., Nishimura, M., Yamagishi, A., Oshima, T. and Kikuchi, H., Complete genome sequence of an aerobic thermotolerant carnobacterium, Sulfolobus tokodaii strain7 DNA Res. 8 (4), 123-140 (2001)
JOURNAL MEDLINE	21456156 11572479
PUBMED	2 (bases 1 to 253050) Director-General of Biotechnology Center, Kawarayabasi, Y., Tanaka, T., Hino, Y. and Kikuchi, H. Direct Submission Submitted (22-DEC-1999) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2-49-10 Mishihara, Shibuya-Ku, Tokyo 151-0066, Japan (E-mail:biomnite.go.jp, Tel:81-3-3481-1933, Fax:81-3-3481-8424) Kawarayabasi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305-0046, Japan Yamagishi, A. and Oshima, T. are at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo 192-0392, Japan The other authors are at the National Institute of Technology and Evaluation, Mishihara 2-49-10, Shibuya-Ku, Tokyo 151-0066, Japan Homology analysis is performed by Smith-Waterman algorithm against EMBL release 109; EMBL release 56.0; Swissprot release 38.0; PIR-Protein release 62.0; and OWL release 31.4.
COMMENT	ORF organization, sequence alignment and more information are available at W.W.W. site of Biotechnology Center, URL: http://www.bio.nite.go.jp/dogan/genome_list-e.html/ Location/Qualifiers 1..253050 /organism="Sulfolobus tokodaii str. 7" /mol_type="genomic DNA" /strain="7" /db_xref="taxon:273063" 480..1019 /gene="ST0846" 480..1019 /codon_start=1 /transl_table=1 /product="179aa long hypothetical protein" /protein_id="BAB56859.1" /db_xref="GI:15621866" /translation="MEKDITYGDEVDYEAVVIESLDLLSLTLMREGTTNSACGA NAVVALDIDHYQNGEPDPFSNYSRKEDVLRLITVMETKKVNTTPPKVEVEII EKIDELKEINDRNGVTL" complement(1500..2816) /gene="STrep24" complement(11500..2816) /note="dispersed repetitive unit, subtype II" /rpt_type="dispersed" /rpt_unit="1500..2816" 1621..2613 /gene="ST0847" 1621..2613 /note="SPIR:S74012 percent identity: 33.935 in 277aa." /codon_start=1 /transl_table=1 /product="330aa long conserved hypothetical protein" /protein_id="BAB65860.1" /db_xref="GI:15621867" /translation="MGIEISGVIPORTLYGMKSKYLDCERILKEIDVTIRREGEKY FTYDYDTLYIKMLIVLVYRESRGITLAANIYVKRFYAKKIPSKSSLHWVHKF SDRIQLLKVRPEPSIEBQPLTHHLLABEFGRDIRLDSFPALPGKSGIETP HIKILLIDARELIDSKRPHEVLASJDEKRRELPSHFPRDVTVEHGSGWRKGKGLM FGCKFAEVSSKTLMWSKVETLVANVSRSBPVPNVTVLADRVFVERGKIYRRKGL
FEATURES	
source	
gene	
CDS	
gene	
repeat_region	
gene	
gene	
CDS	


```
gene
3365..3868
/gene="ST0848"
3365..3868
/gene="ST0848"
/notes="PIR:S74037 percent identity: 43.293 in 164aa."
/codon_start=1
/transl_table=11
/product="167aa long conserved hypothetical protein"
/protein_id="BAB65861.1"
/db_xref="GI:15621868"
/translation="MDTSFLIDWVKYDKRDLFDYTTVFLSESVLNEIRIETPLLMT
SEWLAKRIKILEETVDVRKALLVDVMTREPLSADYPEAVCLVLGKELNDVLTSE
NGVPAKEIIEEYNSVNYRGIDILYIISOKGLISDFINEVKYINSYKHYSEIR
EYKGEI"
4482..5090
/gene="ST0849"
4482..5090
/gene="ST0849"
/notes="PIR:T00006 percent identity: 53.247 in 77aa."
/codon_start=1
/transl_table=11
/product="202aa long hypothetical pepk protein"
/protein_id="BAB65862.1"
/db_xref="GI:15621869"
/translation="MGITKDISKNDIKNAIIEIYKSGKSTPSKOBLELKKRSNTE
IDNIIEDLVKGFIIKIKOGKNKEIOFMFIEGKVDENIGKTEDIEGSDQIKAL
ESAPIVLSKYERLIEDLKIIIPDRIPTEPDLRVYEPFARKSLGHASIKDLRIQLG
LSBERFMRYREYIIONYELIIGSEBGFIKGVMYGIIRKR"
5099..6433
/gene="ST0850"
5099..6433
/gene="ST0850"
/notes="motif=ATP/GTP-binding site motif A (P-loop)"
/codon_start=1
/transl_table=11
/product="444aa long hypothetical protein"
/protein_id="BAB65863.1"
/db_xref="GI:15621870"
/translation="MCNPSYTVARYPDNDNLVIGRVMDSARKSIINNTINDGWT
ITGSPMGKTTILNAVVKDKRDPVFIYLLSLSDSRISKAAMTIRKLEIYVSKV
FNYLSLHKKEIGYKGFYISKDFKNMLHQESSDDARIKJCYREKDIIGLIEFL
NDLKTQSLEKQVILIDETSKSEIIDELHELINKYIPAIILIDEVKVMNALVR
RTENQIDELKSDKDEILKAYCEKYKELSDIVRNALNLADARPKYSEAL
NECKNGAYIDECIKDYLOKYPIDESLIDYSTALEKRIKLEIKIQAQAVIKYI
HERGKDKDEYGVNDIIPETENTYIVGDKVSNHEKIKYIKVKGLEAVQIKDKKV
KRTVNNVNSDPTIIPVNNLSIKKIYDNDDEKRDILVGMRLGIV"
6481..6894
/gene="ST0851"
6481..6894
/gene="ST0851"
/codon_start=1
/transl_table=11
/product="137aa long hypothetical protein"
/protein_id="BAB65864.1"
/db_xref="GI:15621871"
/translation="WYMISRGIVILITPSFYALLELMIDPSQIKTSPGMKEPF
TPTVSLYPRYVYTLFAPSYLASKLPSMETIYWLIGSTLEIDITLWLDVRPYS
WAMFYVYIGIDIVGLMLIRKKIKERKVR"
6894..7523
/gene="ST0852"
6894..7523
/gene="ST0852"
/notes="Swiss_Prot:Q10858 percent identity: 28.289 in
304aa."
/codon_start=1
/transl_table=11
/product="425aa long hypothetical transporter"
/protein_id="BAB65865.1"
/db_xref="GI:15621872"
/translation="HEBKLSQALALGNIGAGITVMAGVSTTANGPALAFL
ITAVVAMSLNNAELASVFPVEGGVYSFALLSGETIGFVGMFVRIGVATISGAT
ALGFSGLITTPSLPSFLYLLAILIIVLIIIDYIGLTAIEIRIVLVVILIGLVI
```

```
PSIALVTSIGIRLQNPSPPLPHGICITGVLTASINAFNPANGENTITLTPSVKGERTV
PAIVLSLISAFIYMLVTFPSVAVDAIGKNFTNSAPLQIVSTIHAPYVFPVNL
ALATITVTLSTIATERTMOMELDNLPKKGGKSLIILIGVMISLIGVNEV
IALVGNFVFSVSGIEAVVRHGRGVSGFPPVQIISVLSGVFIYSLQFO
SILIGFVTLIGLVHSIHKEIMKEGIFSRKK"
8837..9703
/gene="ST0853"
8837..9703
/gene="ST0853"
/notes="PIR:C75043 percent identity: 28.821 in 229aa."
/codon_start=1
/transl_table=11
/product="288aa long hypothetical protein"
/protein_id="BAB65866.1"
/db_xref="GI:15621873"
/translation="MILGKLABELKRGYKIYGYEKPSLONIAVKMLVKNYGFLLK
RISGRIGETIGISLYKKGDEKDEKAPTEGKEGICPIKLPKPLVIDMSLPNLEDE
BEKKTLLQVNLTLVIRKFLMDGALINSNISLGRKRINDIETDNCIYLDPRGD
IIANEELIRTDVFLIIGGIVDKGRLLKATSKLAEKYPCKKVTILRGSIVGPDEN
KIADIIIGVRFGELEKAIIEYQNSDRIARILVDNORGLIEEYVWMLANOKYV
KLIRLGVNAS"
complement(9806..10966)
/gene="ST0854"
complement(9806..10966)
/notes="Swiss_Prot:P19780 percent identity: 35.897 in
117aa."
/codon_start=1
/transl_table=11
/product="386aa long hypothetical transposase"
/protein_id="BAB65867.1"
/db_xref="GI:15621874"
/translation="MTKDFVSKAPADISQSLTPAKGELVIKQKPSVTVKEVEFP
YNNIEIEIKFLEGVNIGLEATGVFYLHKLTEGKFTVVPALTEILGKT
DKLDAQRLVAVMTGVIVKSYIPGEIKELRETLRHRNLVNIQVNEIKTELEIA
GKTIKPFDKGQLEKLAKGELSKEKEBELKELGRLNDAEKLTKQVLELKSJ
EMWKEVDMISIKIPKVIETLSKIPGILTAATVIAEPGDISRPNSKARAYAGF
APRTQSGNSHSGMIRNGKILRGGFYLAASARMLBEPNFERLLIARGSVROAT
CALAGKACIVYHVLKDGYYKGIYVKRLRIPKGRVNVDPVGDALDSLSP"
10218..10748
/gene="ST0855"
10218..10748
/gene="ST0855"
/notes="PIR:F71169 percent identity: 37.313 in 67aa."

Query Match
Best Local Similarity 64.7%; Pred. No. 0.27; Length 253050;
Matches 77; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

gene
CDS
134901 AACCAAGATGAATCTATAATAGCACAGAAAGCAAAATTATATGCGAAAGATATT 134842
QY 517 AATAAGATATTTGACACACGTTGACACAGATTGATGATGCCCAAAAAAGATA 576
Db 134901 AACCAAGATGAATCTATAATAGCACAGAAAGCAAAATTATATGCGAAAGATATT 134842

QY 577 GGATCGGTTGATGTTGCAACGCAATTATATGCTGATGTATATGAAGATTCA 635
Db 134841 GGAGCGGTTTGTATATAGCCTCAGCTGTATGTTCAATAGATTATAGAAAGATTTA 134783

RESULT 23
AC009757/c AC009757 166918 bp DNA linear PRI 27-AUG-2002
LOCUS Homo sapiens chromosome 5, clone RP11-37w16, complete sequence.
DEFINITION AC009757
ACCESSION AC009757
VERSION AC009757.8 GI:8072443
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 166918)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 5, clone RP11-37w16
JOURNAL Unpublished
```



```

Gene
<1..>732
/length="1314"
/gene="emm"
/locus_tag="emm93.1"
<1..>732
/gene="emm"
/codon_start=1
/transl_table=1
/product="M protein precursor"
/protein_id="AAN46658.1"
/db_xref="GI:24286599"
/translation="AVTACGPNQTEYKAEKNTRYNEAHQALREYLGAMNVOQLRI
ALGSHFNDELQKTIQKERELELDQALIKRLINERHHDKREAEKRLAETDLCAD
OEHLDGALRYINKEKAEKREKAEQKLIKKEKQVSAASQGLRDLDAERAKQYVEK
DLANLTAEPRDKYKEKQVSDASRQGLRDLDAERAKQVEKDLANLTAEJDYKYEK
QVSDASRQGLRDLDAERAKQYVEK"
<1..45
/gene="emm"
<46..>732
/gene="emm"
/product="M protein"

```

Query Match	3.7%;	Score 47.8;	DB 1;	Length 732;
Best Local Similarity	46.5%;	Pred. No 3.6;	Mismatches 177;	Indels 0; Gaps 0
Matches 154;	Conservative	0;	Mismatches 177;	Indels 0; Gaps 0
Qy	865	GTGATGACCAAGTTCATATGTC	CCAAATTACAGTTTATGAAAGGAA	TGAGGAAATGCGT 924
Db	196	GAGTTGAAAGACCTTAAAGATGCT	GAGTTGAAGCACTTATGAAGAGAC	ATGATGATCAT 255
Qy	925	GAAAAATACGACTCAGACCCAGAC	CTTATATTTAAAGAGTTAGATCAT	TCGTTTAGGCT 984
Db	256	GACAAAGAGAAAGCAGACCGAAAG	CTTTAGAACAAATTTAGCGATTA	GCAAAACAT 315
Qy	985	TTGACTGTTGCATTAAAGACATGA	AAAAAGCGTTCAAGCATTTAACAA	AAATCAGAG 1044
Db	316	CTTGACGGTGCTTCCTCGTATATCA	TATGAAAAAGAAAGCAGAGGAAA	AGAAAAAGAACGA 375

Oy	1045	GTTCCATTGAACCTGATGTCCAACCAGTGGTGACCGTTGTCAAGATTCCTGCT	1106
Db	376	GACACAGAAAAAATTAAAGAAGAAAAACAAGTCACAGCAGAACCGCTCAAGCTTCGC	435
Oy	1105	TGTGTGGTGGTGTGTTCCAGGTGCTGTGTGATACAGATCAATAGCTTATTAGGTTG	
Db	436	CGTGACTTGAGACGATCACGTGAAGCTTAAGAAAACAGGTGAAAAAGATTACCAACTTG	495
Oy	1165	GAAATTCAGTGGGAAATTTTAAGCAGAAAA	1195
Db	496	ACTGCTGAATTTGATTAAGTTAAAGAAAGAAA	526

RESULT 26	AY139423	923 bp	DNA	linear	BCT 14-NOV-2000
LOCUS	AY139423				
DEFINITION	AY139423				
ACCESSION	AY139423				
VERSION	AY139423.1	GI:24954681			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					

```

gene
CDS
    /gene="emmn"
    /note="emmn98"
    <1..>923
    /gene="emmn"
    /codon_start=1
    /transl_table=1
    /product="M protein"
    /protein_id="AAN64696.1"
    /db_xref="GI:24954682"
    /translation="SAVALITLVGAGIASSTEVKADRYDTHANAVTQGRTPVLDNRLLIL
ENDKSKLRSSENERLQAGLODEKRENEEIOAGIQEERELIDDKDAELKRLNREPHD
LDREERKALDEKLDKQERLDGALRYINKEKAEKKEAKGKLTKEKQISDSRQSL
LRDLDDASREAKKQVEKQDKNALVIAELIDRYKEKQISDSRQSLRDLDDASREAKQVE
KQALNLTVELIDRYKEKQISDSRQSLRDLDDASREAKKQVEKQKALNEANSKLALBETL
NKELEBSKGLTEKKELQAKTAEAKKALV"

```

	Query Match	Best Local Similarity	Score 47.8;	DB 1;	Length 923;
	Matches 160;	Conservative 0;	Mismatches 187;	Indels 0;	Gaps 0;
Qy	849	AGAGAAAGCTCTGTGCTATGACCACTTTAATAGTGCATTTACAGTTATATAGA	908		
Dy	231	ACAGAAAAGAGACAGAGTTGAAAGACCTTAAAGATGCTGAGTTGAAACGACTTAAAGA	290		
Qy	903	ATTGAGGAAAATGCGTAAAAATACGACTCAGCCAGACCTTATATTAAGAGTTAGA	968		

```

Db      291 AGAGAGACATGATCATGACAAAAGAGAGACAGCCAAAAGCTTTAGAGACAAATTAC 350
Qy      969 TCATTCTGTGAGCCTTTGACTGCTGATTAAGAAACAGAGAAAAGGTTCAAGCATT 1028
Db      351 GGATTAAGCAAGAACATCTTGACGTCCTTCCTGCTATATCAATGAAAAAGAGCAGCG 410
Qy      1029 AACACAAAATCAGAGGTTCCATTGAACTGATGTCGAACCCAGTTGTTGACCGTTG 1088
Db      411 AAAAGAAAAAGAGCAGACAGAAAAAACTTAAAGAAAGAAAAACAATCTCAGACGCAAG 470
Qy      1089 TCAAGAGATTCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1148
Db      471 CCGTCAAGGCTTTCGCGCTGACTTGACCCATCAGCTGAAGCTTAAGAAACAAGTTGAAA 530
Qy      1149 AGCTGATTAATGTTGGAAGAAATCAAGTGGAAATTTAAGCAGAAA 1195
Db      531 AGATTGCAAACTTGACTGCTGAATGATTAAGTTAAAGAGAAA 577

RESULT 27
SPERM41      1288 bp      DNA      linear      BCT 12-MAY-1993
DEFINITION   S.pyogenes for emm41 gene.
ACCESSION    X58178
VERSION      X58178.1 GI:47362
KEYWORDS     emm41 gene; M protein; PCR; virulence factor.
SOURCE       Streptococcus pyogenes
ORGANISM     Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
              Streptococcus.
REFERENCE    1 (bases 1 to 1288)
AUTHORS      Podbielski, A.
TITLE        Three different types of organization of the vir regulon in group A
              streptococci
JOURNAL      Mol. Gen. Genet. 237 (1-2), 287-300 (1993)
MEDLINE      93204905
PUBMED       8455563
REFERENCE    2 (bases 1 to 1288)
AUTHORS      Podbielski, A.
TITLE        Direct Submission
              Submitted (28-FEB-1991) A. Podbielski, Institut f Med
              Mikrobiologie, an der Rhein Westf Techn Hochschule, Pauwelsstr
              Klinikum, 5100 Aachen
JOURNAL
COMMENT      See also X58157-X58168 for related sequences.
FEATURES
  source      1..1288
              /organism="Streptococcus pyogenes"
              /mol_type="genomic DNA"
              /strain="C101/103/2"
              /db_xref="taxon:1314"
              17..1256
              /gene="emm41"
              17..23
              /gene="emm41"
              30..1256
              /gene="emm41"
              /codon_start=1
              /transl_table=11
              /product="M protein type 41"
              /protein_id="CAA41167.1"
              /db_xref="GI:47363"
              /db_xref="GOA:O54837"
              /db_xref="TrEMBL:O54837"
              /translation="MIRKDTNRHYSLRKIKITGTASVAVAVALGAFANQTEVKAEN
              ALTAQAEALADIVANTPNQLRDPYAFAPRNNELEKTIQEKRELEGDLAEIKR
              LNEERHDHDKKEARKALEDKLADQEHLDGALRYINKEAEKKEKEAEOQKLKEEK
              ISDASRGRLRLDASREAKQVEKDLANLTAPDPKVEKOVSDASRGQLRDLSDAS
              REAKQVEKDLNLTAPDPKVEKOVSDASRGQLRDLSDASREAKQVEKALEREAS
              KLAALKEKLEBSKULTEKAKELQATLREBAKALQGLAKQAEELAKLAPEKASG
              SOTPAKGTAKVAPKGAQPAQGTGPNQAKPMKETKQLPSTGSAANPFYAAALTV
              MATAGVAAVVRKEEN"
  gene        repeat_region
              507..827

```

```

misc_feature      /note="C repeat"
                  576..632
                  /gene="emm41"
                  /note="spacer 1"
misc_feature      702..758
                  /gene="emm41"
                  /note="spacer 2"

ORIGIN
Query Match      3.7%; Score 47.8; DB 1; Length 1288;
Best Local Similarity 46.5%; Pred. No. 3.4; Indels 0; Gaps 0;
Matches 154; Conservative 0; Mismatches 177;

Qy      865 GTGTATGACACAGCTTAATAGTCCAAATTATACAGTTATGAGAGAAATGCGT 924
Db      300 GAGTATGAAGAGCGCTTAAGATGCTGATGAGAGCACTTAAGAGAGAGATATCAT 359
Qy      925 GAAAAATACGACTCAGACCCAGAGACTTATATTAAGAAGTTAGATCATTTGTTAGCCT 984
Db      360 GACAAAAAAGAGCAGAGCAAGAAAGCTTTAGAGACAAATTAGCGATTAAGCAAGACAT 419
Qy      985 TTGACTGTTGCGATTAAAGACATCAAGAAAGGCTTAACAGCTTAACCAAAAATCAGAG 1044
Db      420 CTGACGGTGCCCTTCGCTATATCAATGAAAAAGAGCAGCGAAAGAAAAAGAGCA 479
Qy      1045 GTTCCAAATGAACCTGATGTCCAACCCAGTTGTTGACCGTTGTCAGAGATTCTCGT 1104
Db      480 GAGCAGAAAAAACTTAAGAGAAAAAACAATCTCAGACGCAAGCCGTCAGAGTCTTCGT 539
Qy      1105 TGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1164
Db      540 CGTGACTTGAGCGCATCACTGTAAGCTTAAGAAACAGTTGAAAAAGATCTACAAACTTG 599
Qy      1165 GAAATCAAGTGGGAAATTTTAAGCAGAAA 1195
Db      600 ACTGCTGAATTTGATTAAGTTAGAGAGAAA 630

RESULT 28
AY139408      949 bp      DNA      linear      BCT 14-NOV-2002
LOCUS        Streptococcus pyogenes isolate NS2660 M protein (emm) gene, emm70
DEFINITION   type, partial cds.
ACCESSION    AY139408
VERSION      AY139408.1 GI:24954582
KEYWORDS     Streptococcus pyogenes
SOURCE       Streptococcus pyogenes
ORGANISM     Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
              Streptococcus.
REFERENCE    1 (bases 1 to 949)
AUTHORS      Dyall-Smith, M.L., Krejany, S., Sriprakash, K.S., Delvecchio, A. and
              McMillan, D.
TITLE        emm sequences of Australian GAS isolates
JOURNAL      Unpublished
AUTHORS      Dyall-Smith, M.L., Krejany, S., Sriprakash, K.S., Delvecchio, A. and
              McMillan, D.
TITLE        Direct Submission
              Submitted (06-AUG-2002) Microbiology and Immunology, University of
              Melbourne, (Gretan Street, Parkville, Victoria 3010, Australia
              Sequence identified in 2002 by students participating in the
              Molecular Microbiology Techniques Course (526-321), Department of
              Microbiology and Immunology, University of Melbourne, Australia
FEATURES
  source      1..949
              /organism="Streptococcus pyogenes"
              /mol_type="genomic DNA"
              /isolate="NS2660"
              /db_xref="taxon:1314"
              /country="Australia: Northern Territory"
              <1..>949
              /gene="emm"

```

CDS

```
/note="em70"  
<1..>949  
/gene="em70"  
/codon_start=3  
/transl_table=11  
/product="M protein"  
/protein_id="AAN64681.1"  
/db_xref="GI:24954583"  
/translation="ASVAVLVLAGFANOTEVKAEHESVTRAREAAIREMROGR  
RINERHDHREARERKLEKLDKQSHNGALRYINERKEKEKEKQKLEK  
QISDASROGRLRIDASREAKQVEKDLANTLAEIDYKEKQISDASROGRLRID  
SREAKKVEADLALTAHOKLEKQISDASROGRLRIDASREAKKVEADLAEAN  
SKLOALEKNEKEBEKKSSEKKEALQRLAEKALKEQ"
```

ORIGIN

```
Query Match      3.6%; Score 46.4; DB 1; Length 949;  
Best Local Similarity 46.2%; Pred. No. 7.3; Matches 192; Conservative 0; Mismatches 221; Indels 3; Gaps 1;
```

```
865 GTGTATGACGAGCTTAATAGCCAAATTTACAGTTATGAAGAAATTTAGGAAATGGCT 924  
276 GAGTTAAAGACCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 335  
925 GAAATAATCACTCAGACCCAGAGACTTATATTAAGATTGATGATGATGATGATGATGATGAT 395  
336 GACAAAAGAGACGAGCGAAGGTTTGAAGACAAATTAAGCGATTAAGCAAGAAACAT 395  
985 TTGACTGTTCGATTAAGAAATCAGAAAAGGTTTCAAGCATTAACACAAAATTCAGAG 1044  
396 CTTAAGCGTCCCTTGCCTATATCAATGAAAAAGAGCAGCGCAAGGTTCAAGGCTTCGT 1104  
1045 GTTCCAAATTAAGCCTGATGATCCAAACCCAGTTGTTGACCGTTGCAAGATTCCTGAT 1104  
456 GAGCAGAAAAAATTAAGAAAGAAAAAACAATTCAGACGCAACCGTTCAGAGGCTTCGT 1104  
1105 TGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1164  
516 CCGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1164  
1165 GAAATAATCAAGTGGAAATTTTAAAGCAAGAACTTTGAAATTCAGATTTTCAATAT 1224  
576 ACTGCTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1224  
1225 GTTAACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1280  
633 GGTCTTGGCGGTGATCTGTGATCATCAAGCAAGTAAAGAAAAAGTAGAAGCAGA 688
```

RESULT 29
AY139421

LOCUS AY139421 951 bp DNA linear BCT 14-NOV-2002
DEFINITION Streptococcus pyogenes isolate NS80 M protein (emm) gene, emm70
ACCESSION AY139421
VERSION AY139421.1 GI:24954669
KEYWORDS Streptococcus pyogenes
SOURCE Streptococcus pyogenes
ORGANISM Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus
1 (bases 1 to 951)

REFERENCE

1 Dyal1-Smith, M.L., Krejany, S., Striprakash, K.S., Delvecchio, A. and
Mcmillan, D.
emm sequences of Australian GAS isolates
2 (bases 1 to 951)
Unpublished
Dyal1-Smith, M.L., Krejany, S., Striprakash, K.S., Delvecchio, A. and
Mcmillan, D.
Direct Submission
Submitted (06-AUG-2002) Microbiology and Immunology, University of
Melbourne, Grattan Street, Parkville, Victoria 3010, Australia
Sequence identified in 2002 by students participating in the

FEATURES
source

Molecular Microbiology Techniques Course (526-321), Department of
Microbiology and Immunology, University of Melbourne, Australia
1..951
/organism="Streptococcus pyogenes"
/mol_type="genomic DNA"
/isolate="NS80"
/db_xref="taxon:1314"
/country="Australia: Northern Territory"
/gene="em70"
/note="em70"
<1..>951

CDS

ORIGIN

```
Query Match      3.6%; Score 46.4; DB 1; Length 951;  
Best Local Similarity 46.2%; Pred. No. 7.3; Matches 192; Conservative 0; Mismatches 221; Indels 3; Gaps 1;
```

```
865 GTGTATGACGAGCTTAATAGCCAAATTTACAGTTATGAAGAAATTTAGGAAATGGCT 924  
274 GAGTTAAAGACCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 333  
925 GAAATAATCACTCAGACCCAGAGACTTATATTAAGATTGATGATGATGATGATGATGATGAT 395  
334 GACAAAAGAGACGAGCGAAGGTTTGAAGACAAATTAAGCGATTAAGCAAGAAACAT 395  
985 TTGACTGTTCGATTAAGAAATCAGAAAAGGTTTCAAGCATTAACACAAAATTCAGAG 1044  
394 CTTAAGCGTCCCTTGCCTATATCAATGAAAAAGAGCAGCGCAAGGTTCAAGGCTTCGT 1104  
1045 GTTCCAAATTAAGCCTGATGATCCAAACCCAGTTGTTGACCGTTGCAAGATTCCTGAT 1104  
454 GAGCAGAAAAAATTAAGAAAGAAAAAACAATTCAGACGCAACCGTTCAGAGGCTTCGT 1164  
1105 TGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1164  
514 CCGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1224  
1165 GAAATAATCAAGTGGAAATTTTAAAGCAAGAACTTTGAAATTCAGATTTTCAATAT 1224  
574 ACTGCTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1280  
1225 GTTAACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1280  
631 GGTCTTGGCGGTGATCTGTGATCATCAAGCAAGTAAAGAAAAAGTAGAAGCAGA 686
```

RESULT 30
AX083744/c

LOCUS AX083744 1141 bp DNA linear PAT 28-FEB-2001
DEFINITION Sequence 22 from Patent W00110161.
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 Kunet, L. and Clemens, S.
TITLE Regulation of embryonic transcription in plants

NOTE: Estimated insert size may differ from sequence length (see http://www.hpsc.bcm.tmc.edu/docs/genbank_draft_data.html)

NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will

```

* 1 118527: contig of 118527 bp in length
* 118528 118527: gap of unknown length
* 118628 233719: contig of 115092 bp in length
Location/Qualifiers
1. .233719

```

ORIGIN

Query Match	3.5%	Score 45.8;	DB 2;	Length 233719;
-------------	------	-------------	-------	----------------

Matches 131; Conservative 0; Mismatches 142; Indels 0; Gaps 0

166 CCCCAATTGCAACGGAATGGGAATATCACATATCATCAATACAGAGAAGCCAGA 225

172462 CCCACATTGGAAGTGAGACTTTAAATAATTTGGGATACTATTTACAAAGCAGATTTGA 172521

226 GAAGTTCAGTCACGCATAATCCATTTTTAGAGGCCAACTATATTCATCGTTTACCTTAT 285

172522 AGAGTAACTTTAGGATAAANCACTTTGTAACAGTTTTCATGTTTACIGTTT 172581

286 A I C A C C G A C G A G C A I T T G A I C T I G A A A I C A I C A I T T A C I C A G A C C C I G G A I A I C A I 345

[illegible]

230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

1. The first group of variables includes the demographic characteristics of the respondents, such as age, gender, and education level. These variables are used to control for potential confounding factors that may influence the relationship between the independent and dependent variables.

RESULT 33

Accession	Length	Topology	Accession	Length	Topology
AC125664	243143 bp	linear	HTG 10-MAY-2003		

unordered pieces.
AC13E664
CCE66ON

KEYWORDS: HTG: HTGS PHASE1: HTGS DRAFT: HTGS FIT:TOP
AC125664.3 GI:30522560
EXTENSION

КОРНЕ КАСКУС НОРВЕГИКУС (NORWAY LAL)

REFERENCE
AUTHORS

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1^a (bases 1 to 243143)
Muzny, D., Marie, M., Metker, M., Lee, S., Adams, C., Alder, J.,
Barnes, R., Breen, J., Brown, L., Burdette, D., Busch, W., et al.

Allylen, C., Allen, H., Alldrick, S., Mullin, R., Arguiano, J.,
Anyalelechi, V., Aoyagi, A., Ayodeji, M., Bace, E., Baden, H.,
Baldwin, D., Banderianale, D., Barber, M., Barnesead, M., Benahmed, F.,
Biswal, O., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burck, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, J., Ceasear, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
David, M. L., Davis, C., Davy-Carrillo, L., De Anda, C., Dedertich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Duplin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falle, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, B., Foster, P.,
Frazer, C. M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Gazza, M.,
Guebragosorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevarz, W.,
Gumbare, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Hulxy, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karapathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kows, C., Kraft, C. L., Lebow, H., Levant, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuber, L., Louisged, H., Lorado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mamoud, M., Malloy, K., Mangun, A.,
Mangus, B., Mapua, P., Martin, K., Martin, R., Matlitz, E.,
Mawliny, S., McLeod, M. P., McNell, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Mnij, E., Montemayor, J., Moore, S.,
Morjan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, A., Pal, S., Parks, K.,
Nwackeemelun, O., Okwunon, G., Olampungasoon, A., Pal, S., Parks, K.,
Pasternack, S., Paul, H., Perez, A., Perez, L., Pflannkoch, C.,
Piopler, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzao, M., Quiroz, J., Ren, Y., Reuther, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuther, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Saversy, G., Scherer, S., Scott, G., Shademan, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Slason, I., Slinger, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Soreiller, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, T.,
Taylor, T., Thomas, N., Thomas S., Tingey, A., Trejos, Z., Usmani, K.,
Vals, R., Vetz, V., Villasaana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willison, R., Wlecyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weise, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE
JOURNAL
REFERENCE

AUTHORS
TITLE
JOURNAL

REFERENCE

TITLE

COMMENT

On May 10, 2003 this sequence version replaced g1:23269430. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas

Qy	612	TCTGATCTGTAATAAGAAATTTCCAGCCAGCTTGATTAATGACGGTTCAGGTTCTGACA	671
Db	416	KMDMDMBSTYNNNNNGRTTYTGWTGKXKMGTYTYKKKANKCKBRADHKCTHNTTMMGM	475
Qy	672	AAGTGATCCTGAGAAATTCCCCACAGAGTTGAAAAAATTGATTGAAAGTAAGTACGGAAAT	731
Db	476	KTYMNNCKWKSNTGKSHRBAAVYTYMMWRRAHANNNNMDYMKXCTWYKYBVSX	535
Qy	732	CAACATGAAAGATGATCATTCACATACGGAATCAAGTATTAATGCGGAGTCGAAGG	791
Db	536	WMNNYAAAYTSSMWTYSRYRYMKTKNNKWRNRSDTRSMGRANNYRABHYGYKNTTRMB	595
Qy	792	TGGCTCAGAAACCCCAAATTGGTATCAGCACTACTCCAATGAAAAAGAAAAAGCCAGA	851
Db	596	WSHTBHBRAAGAAHYMMBMVBAKCHCMKAYKAKXYAGAGSNNNNNNNNNNNNNNNA	655
Qy	852	AGAAAGCTGTGTGTGTATGACCACTTAATAGTGCCAAATTACAGTTATGAAGAAT	911
Db	656	TCARDYTAASGMYMANAKMTYYBAANNAAYTTAANNMGCMNATDRTRTMMNNNNN	715
Qy	912	GAGGAAATGCGTGAATAATACGACTCAGCCACAGACTTATATTAAGAAGTTAGATCA	971
Db	716	NAGTKNNNNNNNAKVASAAKYAAAVKAAKGMWRANKMARGMHDAABATDCKUNG	775
Qy	972	TTCTGTGAGCCTTGAAGCTGTGGCATTAAGAAKATCAGAAAAAGGTTACAAGCATTAAC	1033
Db	776	AATKTTNNNNNTYGVVNTTAARGMANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	835
Qy	1032	ACAAAAATCAGAGTTCACATGGAACGTGATCCAAACCCAGTGTGGAACGGTTGTCA	1091
Db	836	NNNNNNNNNAAYAMWNTKMYTTTDDMRBAVYTNNNNNNNNAAYGAVADYAYMSPTCAAM	895
Qy	1092	AGAGATTCTGCTGTGTTGGTGGTGTGCT	1121
Db	896	MKWATYKNNATTYNRTAMRTNNNNNNNT	925

```

RESULT 35
AY279126
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source
1..581
/organism="Arisaema harmandi"
/organelle="plastid:chloroplast"
/mol_type="genomic DNA"
/specimen_voucher="W. Heteroscheide H.A.R. 220"
/db_xref="taxon:228818"
<1..>581
/gene="trnL"
<1..>581
/gene="trnL"
/notes="contains intron and flanking exons"
ORIGIN
Arisaema harmandi chloroplast trnL-leu (trnL) gene, partial
sequence.
AY279126
AY279126.1 GI:30983968
chloroplast Arisaema harmandi
Arisaema harmandi
Euarctocata Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Arisaema.
1 (bases 1 to 581)
Renner,S.S., Zhang,L.-B. and Murata,J.
A chloroplast phylogeny of Arisaema (Araceae) illustrates Tertiary
floristic links between Asia, North America, and East Africa
Am. J. Bot. 91 (6), 881-888 (2004)
2 (bases 1 to 581)
Renner,S.S., Zhang,L., Guegan,G. and Murata,J.
Direct Submission
Submitted (18-Apr-2003) Biology, University of Missouri-St. Louis,
8001 Natural Bridge Rd., St. Louis, MO 63121, USA
Location/Qualifiers
1..581
/organism="Arisaema harmandi"
/organelle="plastid:chloroplast"
/mol_type="genomic DNA"
/specimen_voucher="W. Heteroscheide H.A.R. 220"
/db_xref="taxon:228818"
<1..>581
/gene="trnL"
<1..>581
/gene="trnL"
/notes="contains intron and flanking exons"
misc_feature

```

	Query Match	Similarity	3.5% Best Local	Score 45.4	DB 8	Length 581	
	Matches	97	Conservative	0	Mismatches	86	Indels
							Gaps
							0
Qy	878	TTAATAGTCCAAATTTCAGCTTATGAAGAAATTGAGGAAATCGTGAATAATGCACT	937				
Db	352	TAAAAAATTAAATTTTATGAAAAATGAAAGGTTATTCGAAATTCATTTAAATTGGAAAT	411				
Oy	938	CAGACCCAGAGACTTATATTAAAGAGTTAGATCATTCGTGTGAGCCTTTGACTGTTCGA	997				
Db	412	TGAAGTAAAGGTGCAATATTCAGTGATCCAAATCATTTACTCCAGAGTCTGATGATCTTT	471				
Oy	998	TTAAGAATCATGAAAAAGGTTTACACAGCATTTAACACAAAATTCAGAGTTCCAATTGAAC	1057				
Db	472	TGAAAAAATAAAAAAAAAAATGATTAATCCGACAGCAAGAAATTAAGAGAGATCCCATCTCAC	531				
Oy	1058	CTG 1060					
Db	532	ATG 534					

RESULT 36
AC0008121/c

LOCUS	105989 bp	DNA	linear	PRI 01-MAR-2003	
DEFINITION	AC0008121	Homo sapiens 12 BAC RPL11-407N8 (Rosewell Park Cancer Institute Human BAC library) complete sequence.			
ACCESSION	AC0008121				
VERSION	AC0008121.43	G1:26626577			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 105989)				
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrook,S.L., Amaratunge,H.C., Are,J.R., Ayale,M., Banks,T., Barberia,J., Benton,J., Bimsge,K., Blankenburg,K., Bonini,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bahay,C., Burch,P., Burkett,C., Burnett,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davis,M.L., Davis,C., Davy-Carroll,L., Dedrich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Doughwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Barnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escoto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gali,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,V., Hernandez,O., Hodgson,A., Hognes,M., Hollway,C., Hollins,B., Homsl,P., Howard,S., Huber,J., Huliyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kurehi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Liu,C., Liu,J., Liu,W., Louisged,M., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Merdonel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., McWhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzger,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabdat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkw.S., Oguh,M., Okumou,G., Orangunye,N., Ovielo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Prims,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojudoakan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shendhan,I., Shm,C., Shooshitari,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svarek,A., Tabori,P., Tameris,A., Tameris,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thoms,S., Uমান,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,				

Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Kuchera-Lapate, R., Weinstein, G., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 105989)
Worley, K. C.
Direct Submission
Submitted (24-JUL-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 105989)
Worley, K. C.
Direct Submission
Submitted (27-FEB-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 105989)
Worley, K. C.
Direct Submission
Submitted (01-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 1, 2003 this sequence version replaced g1:28570287.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

SOURCE

Location/Qualifiers
1..105989

/organism="Homo sapiens"
/mol_type="genomic DNA"

/db_xref="taxon:9606"
/chromosome="12"

/clone="RP11-407N8"
1..2004

/note="overlaps bases 194739..196742 of clone AC087884"
/function="clone overlap"

repeat_region
2205..2312
/rpt_family="MER99"

repeat_region 3552..3836
/rpt_family="AluSx"
repeat_region 3837..3859
/rpt_family="TAAAA)n"
repeat_region 3883..3903
/rpt_family="AT_rich"
STS 4027..4118
/standard_name="STS-241143"
repeat_region complement(5005..5313)
/rpt_family="AluSg"
repeat_region complement(5410..5718)
/rpt_family="AluSx"
repeat_region 5756..5889
/rpt_family="AluJo"
repeat_region 5890..6177
/rpt_family="AluSx"
repeat_region 6178..6325
/rpt_family="AluJo"
repeat_region 6385..6584
/rpt_family="MLT21"
repeat_region 6585..6627
/rpt_family="(TCC)n"
repeat_region complement(6628..6888)
/rpt_family="AluJo"
STS 6669..6770
/standard_name="D11S2921"
repeat_region 6900..6959
/rpt_family="MLT21"
repeat_region 7157..7282
/rpt_family="AluJo/FLAM"
repeat_region 7285..7592
/rpt_family="AluJo"
repeat_region complement(7594..7860)
/rpt_family="AluSx"
repeat_region 8059..8129
/rpt_family="HERVL"
repeat_region 8130..8353
/rpt_family="MER20"
repeat_region 8354..9308
/rpt_family="HERVL"
repeat_region 9309..9619
/rpt_family="AluJo"
repeat_region 9620..9709
/rpt_family="HERVL"
repeat_region 9710..9827
/rpt_family="FLAM_C"
repeat_region 9838..10226
/rpt_family="HERVL"
repeat_region 10251..10871
/rpt_family="HERVL"
repeat_region 10873..11332
/rpt_family="MLT21"
repeat_region complement(11333..11643)
/rpt_family="AluSx"
repeat_region 11644..11804
/rpt_family="MLT21"
repeat_region 11843..12017
/rpt_family="L2"
repeat_region complement(12018..12313)
/rpt_family="AluSx"
repeat_region 12314..12473
/rpt_family="L2"
repeat_region complement(12314..12473)
/rpt_family="AluSg"
repeat_region 12474..12498
/rpt_family="(TG)n"
repeat_region 12499..12794
/rpt_family="L2"
repeat_region complement(12499..12627)
/rpt_family="AluSg"
repeat_region 13789..14044
/rpt_family="AluJo"
repeat_region 14782..15087

repeat_region /rpt_family="AluSx"
15312..15611
/rpt_family="AluSp"
repeat_region complement(15692..15996)

Query Match 3.5%; Score 45.4; DB 9; Length 105989;

Best Local Similarity 48.0%; Pred. No. 8;

Matches 130; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

624 TGAAGATTTCAGCCAGCTTTGTAATGACGTTCTGTAAGATGATCTGA 683

92156 TAACTATAGCAAGAGTCTTTTAAATGTTATGCTGTAATTTAGTACATAA 92097

684 GAAGTTCACCAAGAGTTGAAAAATTTGATTTGAACTGGAATTCATGAAAG 743

92096 CGAATTAATTCCTTTGTAATTAACCAACTGAAATTAATTAATGACCTAATGATAA 92037

744 ATGTACATTACCAATCGAATCAAGTTTAAATGCTGACGTCAGAGGCTCAGAAAC 803

92036 TGTGTAATTAACCAAACTGACCAAGAAAGAAATGAAATCTCTATAGCTTATAGATG 91977

804 ACCCAATTTGTTATCGAGTACTCCATGAAAGAAAGCAAGAAAGCTCTGT 863

91976 TAAAGGAATTAATGATGATTTCTTTTCAAAAATTAATAACCAAGCTCAGAACAA 91917

864 TGTGTAATGACAGCTTAATATGCTCAATTTA 894

91916 TGAATTTCTACCAATTAACCAAGACAGATCA 91886

RESULT 37
AC009727.2
WPCOMMENT

Sequence split into 4 fragments LOCUS AC009727 Accession AC009727

Fragment Name Begin End

AC009727_0 1 110000

AC009727_1 100001 210000

AC009727_2 200001 310000

AC009727_3 300001 374634

Continuation (3 of 4) of AC009727 from base 200001 (AC009727 Homo sapiens chromosome 12

Query Match 3.5%; Score 45.4; DB 2; Length 110000;

Best Local Similarity 48.0%; Pred. No. 8;

Matches 130; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

624 TGAAGATTTCAGCCAGCTTTGTAATGACGTTCTGTAAGATGATCTGA 683

70307 TAACTATAGCAAGAGTCTTTTAAATGTTATGCTGTAATTTAGTACATAA 70366

684 GAAGTTCACCAAGAGTTGAAAAATTTGATTTGAACTGGAATTCATGAAAG 743

70367 CGAATTAATTCCTTTGTAATTAACCAACTGAAATTAATTAATGACCTAATGATAA 70426

744 ATGTACATTACCAATCGAATCAAGTTTAAATGCTGACGTCAGAGGCTCAGAAAC 803

70427 TGTGTAATTAACCAAACTGACCAAGAAAGAAATGAAATCTCTATAGCTTATAGATG 70486

804 ACCCAATTTGTTATCGAGTACTCCATGAAAGAAAGCAAGAAAGCTCTGT 863

70487 TAAAGGAATTAATGATGATTTCTTTTCAAAAATTAATAACCAAGCTCAGAACAA 70546

864 TGTGTAATGACAGCTTAATATGCTCAATTTA 894

70547 TGAATTTCTACCAATTAACCAAGACAGATCA 70577

SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS
Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katsigiri, F., Quan, S., Tao, Y., Whitman, S., Xie, Z., Zhu, T. and Zou, G.
TITLE
Plant genes involved in defense against pathogens
JOURNAL
Parent: WO 0300898-A 5263 03-JAN-2003;
Syngenta Participations AG (CH)

FEATURES
source
1..2000
Location/Qualifiers
/organism="Oryza sativa"
/mol_type="unassigned dna"
/db_xref="taxon:4530"

ORIGIN

Query Match 3.5%; Score 45; DB 6; Length 2000;

Best Local Similarity 8.4%; Pred. No. 14;

Matches 53; Conservative 290; Mismatches 285; Indels 1; Gaps 1;

189 GGAATATCATCATATCAATCAATACAGAGAGCCAGAGAGTTCAGTCAAGTAAATCC 248

383 SRRKRKCMRKRKYKRMGYSRMSCKRARMKCRSRARWRCRCRCMTKMSYGM 442

249 ATTTTAGAGGACATATATCATGCTTTTACCTTATATCAACGAGCAGGACATTTGA 308

443 RWSWRWASKTKMWSRMRKRKCKSRITTMGKTRGCMGTGRCRYKRGRMKCR 502

309 TCTGAAATCATCATTTACTCAGACCTGATATCATCAAGAGATGATGAAACCA 368

503 RRMGRWYRWRWYRYSARITWRYCARKKYSARAKRCWYRGXYWAGMMKRYKM 562

369 GACATCCTCGAATGAGAGAAAAACATTTCTTACATTTCTGCTGACATTAACGAATGA 428

563 YYYKMMWYKRYKSCSWYCKMSYASCKMSRKAGAKCKRSKMSMSKMSRSKCRK 622

429 AAGACCGGATTAAGTTCATCGGAGATTAAGTTCAGTTGTCACAGGTTTATTC 488

623 CASKSSAKRYAMMGMTSGSRMSRKSTCYWRKMSKSTCTMYYSKITYAKYSY 682

489 CCATTTTATCCCAATGTTATCAGATGATGATTAAGATATTTGCAACAGTTCACAGAT 548

683 WRYRRAWCMYWRWYRYSYRYMYTSTMAATGMKYSGRYWTWYKCKSMK 742

549 TGCACATGG-TTATGCCCAAAAAAGATGATGATCTGGGTTGATGTTGCACTGCAATTT 607

743 YRSMMYYSMMWAKTWKMRRYATRMWMMYYSMMKWTCTMGWYMYWRTYMKRM 802

608 ATGCTCATGTTTATATGAAATTTCAAGCTTTGATTAATGACGTTTCAGGTTTC 667

803 WYKCTTYWYSATYWTGTWAAWMAKTMWRMGATKRGARARWMAWMAWMAWMAW 862

668 TGAAGATGATCTGAGAGTTCCCAACAGAGTTGAAAAATTTGATTTGAACTGAG 727

863 WMTKKGAKMAWTMAKMKRYKYSWMAWYRYTRRYRTTCWMAWMAWMAWMAWMAW 922

728 AATTCAACATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787

923 KGSAXMMWMMWKGKRGWMTKYWYCTTWMAAGRAATKWCAGWMAWMSYSWTRTYMRT 982

788 AGGGTGGCTCAGAAACACCAATTTGTA 816

983 WRMMASSRTAKRMARMMWKTAWMSKYA 1011

RESULT 39
AR450463
LOCUS
DEFINITION
Sequence 1127 from patent US 6673910.
ACCESSION
AR450463
VERSION
AR450463.1 GI:42679520

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 04:40:39 ; Search time 3059 Seconds
(without alignments)
15474.080 Million cell updates/sec

Title: US-10-069-062-6
1299

Perfect score: 1 atgcacaaagcattgctgc.....aagactatagattcataa 1299

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	150.6	11.6	1059	CNS06MRJ	AL405845 T3 end of
C 2	131.8	10.1	920	CNS06MRJ	AL419103 T3 end of
C 3	71.8	5.5	973	CNS07ANQ	AL417100 T7 end of
C 4	65.6	5.1	917	CNS06MRJ	AL418769 T3 end of
C 5	63.8	4.9	920	CF824043	CF824043 EST701425
C 6	62.2	4.8	846	CO014330	CO014330 EST802665
C 7	62.2	4.8	922	CO008271	CO008271 EST796606
C 8	60.6	4.7	488	AZ931444	AZ931444 474_dh284
C 9	60	4.6	482	BZ304881	BZ304881 KD3193.P1
C 10	55	4.2	466	AU268449	AU268449 AU268449
C 11	55	4.2	479	AU268448	AU268448 AU268448
C 12	53.4	4.1	717	BJ343297	BJ343297 BJ343297
C 13	53.4	4.1	1101	CNS0039G	AL063921 Drosophila
C 14	51.6	4.0	322	AU052799	AU052799 AU052799
C 15	49.2	3.8	360	CO142585	CO142585 EST837226
C 16	49.2	3.8	412	CO14615	CO14615 EST829286
C 17	49.2	3.8	412	CO142023	CO142023 EST836694
C 18	49.2	3.8	461	CO148810	CO148810 EST838663
C 19	47.8	3.7	775	B0862174	B0862174 OGC20P15.
C 20	47.8	3.7	775	B0862174	B0862174 OGC20P15.
C 21	47.6	3.7	447	CO152593	CO152593 EST827646
C 22	47.6	3.7	475	CO137286	CO137286 EST831967
C 23	47.6	3.7	478	CO148500	CO148500 EST823553
C 24	47.6	3.7	561	CO136871	CO136871 EST831542

25	47.6	3.7	803	CO014331	CO014331 EST802666
26	47.6	3.7	1101	CNS0106X	AL098595 Drosophila
27	46.6	3.6	1101	CNS0182P	AL108811 Drosophila
28	46.2	3.6	716	B0866497	B0866497 OGC8C07.Y
29	46.2	3.6	745	B0870439	B0870439 OGD9A23.Y
C 30	46.2	3.6	1101	CNS017KX	AL108171 Drosophila
C 31	45.8	3.5	727	CG067021	CG067021 PUIDB42TD
C 32	45.4	3.5	499	CNS06D75	AL393447 T3 end of
C 33	45.2	3.5	715	B0860948	B0860948 OGC17C09.
C 34	45	3.5	697	BM165227	BM165227 EST567730
C 35	45	3.5	780	BM169836	BM169836 EST572359
C 36	45	3.5	805	BM161967	BM161967 EST564490
C 37	44.8	3.4	486	AQ400544	AQ400544 HS_5064.A
C 38	44.6	3.4	1101	CNS00DGI	AL069971 Drosophila
C 39	44.6	3.4	900	AZ549980	AZ549980 ENTDP94TF
C 40	44.6	3.4	1101	CNS00DGI	AL069940 Drosophila
C 41	44.4	3.4	384	CO224433	CO224433 MS01021.B
C 42	44.4	3.4	705	B0869724	B0869724 QGD7A12.Y
C 43	44.4	3.4	802	CO122187	CO122187 GR_Eb03L
C 44	44.2	3.4	637	CR331656	CR331656 Medicago
C 45	44.2	3.4	752	BG585318	BG585318 EST487082

ALIGNMENTS

RESULT 1
CNS06MRJ/c
LOCUS
DEFINITION
T3 end of clone AU0AA006F09 of library AU0AA from strain CBS 3082
ACCESSION
AL405845
VERSION
AL405845.1 GI:12169063
KEYWORDS
SOURCE
ORGANISM
Saccharomyces kluyveri
Saccharomyces kluyveri
Saccharomyces kluyveri
Saccharomyces kluyveri
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE
AUTHORS
Soulier, J.L., Aigle, M., Artiguenave, F., Blandin, G., Bolotin-Fukuhara, M., Bon, E., Brotier, P., Casaregola, S., de-Montigny, V., Dujon, B., Durand, P., Lepingle, A., Llorente, B., Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S., Saurin, M., Tekala, F., Toffano-Nicolas, C., Weislowski-Louvel, M., Wincker, P. and Weissenbach, J.

TITLE
Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
JOURNAL
MEDLINE
PUBMED
20584711
FEB8 Lett. 487 (1), 3-12 (2000)

TITLE
Neuvéglise, C., Bon, E., Lepingle, A., Wincker, P., Artiguenave, F., Galliard, C. and Casaregola, S.
Genomic exploration of the hemiascomycetous yeasts: 9.
Saccharomyces kluyveri
FEB8 Lett. 487 (1), 56-60 (2000)

JOURNAL
MEDLINE
PUBMED
11152884
3 (bases 1 to 1059)
REFERENCE
Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: segre@genoscope.cns.fr - Web: www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactic var. lactic, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia strobiliphila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of

COMMENT


```

Db      1 CTGGGTACCATCTAGATGACACGACGACCGGTGTCTCAAGCAACGGCAACGGCAAT 60
Qy      395 TTCTTACCATCTCGTGGCATTACCGAAGTGGAAGAACCGGATTAGTTCACTGGCAG 454
Db      61 TCTTATACCATTAAGAAAGAAATCAACAGAGTTGCCAAAACAGACTTGGGTGCTGGCCG 120
Qy      455 GATTAGTGTCACTGTTGGCAGAGTTTATATCCCATTTTATCCCAA---TGTATCA 511
Db      121 GATTGTTGTCAGTAGTAACGGCGGCACTGCTCTTACTTACTGCAACCGGAGCTTG 180
Qy      512 GTACCAATTAAGATATTTTGGACACAGTTGACAGATTTGACATTTGTTATGCCCAAAA 571
Db      181 ATTCTACCTCCGCAACCGGTCACTAATGACGCAAGTACTCACTGTCAACGCCCAAAA 240
Qy      572 ACATAGGATCTGGGTTGATGTTGCACTGCAATTTATGTTGATTTATATAGAAAT 631
Db      241 AATCGGCTGCTGTTTCAATGTTGACAGCTGTGTACGCTCTTATTTGTTACCCGCTT 300
Qy      632 TTCAAGCAGCTTTGA-----TAAATGACGTGTTTCAAGTTCTAGAAAGTATC 679
Db      301 TCCAGCTAGCTTGAATCGAAGAAATCTAAGACAGAGTTTCAAGTATGCAAGTGACT 360
Qy      680 CTGAG-----AAGTCCCAACAGAGTTGAAAATTTGATGAAAGTACTGGGAAT 730
Db      361 CTACGCTAGCGCTGTGACTCAAGAACTCTTCAAGACATAGTGGGTTCTGACTGGCAT 420
Qy      731 TCAACATGAAGATGTATACATTAAGCAATGCAAGTATTAATAGGTTGACGTTCAAG 790
Db      421 TCAAAATGAAGAAATGTGCTTGGCGCGCGGCTGACACTGTAATGGAAGAGTCTGG 480
Qy      791 GTGGCTCAGAAACACCCAAATTTGTATACAGAGTACTCCATGCAAGAAAGAAAGCCAG 850
Db      481 TCGGTAGGAGACGCGCAATTTGTATGCTATGCTCAAGTGAAGAGAAACCGG 540
Qy      851 AAGAAAGCTCTGTTGTATGACACAGCTTAATATGTCCTCAATTTACAGTTTATGAAG 910
Db      541 TCGAGGGCCAGAACTATTTGAGAGCTTAATAAGCCCAATGAGTGGCTCATCGCGCAT 600
Qy      911 TGAGGGAATGCGTGAATAATGCACTGACGCCAGAGCACTTATA 955
Db      601 TATCTACTGCTTACGAACTCCAGCGGCGCAACGAGCGGCTTACA 645

RESULT 3
LOCUS   CENS07AVQ/c          973 bp    DNA          linear    GSS 08-JUL-2001
DEFINITION
T7 end of clone BC0AA007C01 of library BC0AA from strain CBS 767 of
Debaryomyces hansenii, genomic survey sequence.
ACCESSION
AL437100
VERSION
AL437100.1 GI:12220513
KEYWORDS
GSS.
SOURCE
Debaryomyces hansenii (anamorph: Candida famata)
ORGANISM
Debaryomyces hansenii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Debaryomyces.
REFERENCE
1 (bases 1 to 973)
Soulciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Sautin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Winkler,P. and Weissbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL
FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE
11152876
PUBMED
20584724
AUTHORS
Lepingle,A., Casaregola,S., Neuvéglise,C., Bon,E., Nguyen,H.,
Artiguenave,F., Winkler,P. and Gallardin,C.
Genomic exploration of the hemiascomycetous yeasts: 14.
Debaryomyces hansenii var. hansenii
FEBS Lett. 487 (1), 82-86 (2000)

```

```

MEDLINE
20584724
PUBMED
11152889
REFERENCE
3. (bases 1 to 973)
AUTHORS
Soulciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Sautin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Winkler,P. and Weissbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL
FEBS Lett. 487 (1), 3-12 (2000)
COMMENT
Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxi,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
location/Qualifiers
1..973
/organism="Debaryomyces hansenii"
/mol_type="genomic DNA"
/strain="CBS 767"
/variant="hansenii"
/db_xref="taxon:4959"
/clone="BC0AA007C01"
/clone_lib="BC0AA"
/note="end : 77"
complement(32..>139)
/note="similar to Saccharomyces cerevisiae ORF YMR220w [
BRG8 : phosphomevalonate kinase ]"
/evidence=not_experimental
<278..>972
/note="similar to Saccharomyces cerevisiae ORF YBR166c [
TYR1 : prephenate dehydrogenase (NADP+)]"
1 putative frameshift(s)"
/evidence=not_experimental

misc_feature
complement(32..>139)
/note="similar to Saccharomyces cerevisiae ORF YMR220w [
BRG8 : phosphomevalonate kinase ]"
/evidence=not_experimental
<278..>972
/note="similar to Saccharomyces cerevisiae ORF YBR166c [
TYR1 : prephenate dehydrogenase (NADP+)]"
1 putative frameshift(s)"
/evidence=not_experimental

misc_feature
complement(32..>139)
/note="similar to Saccharomyces cerevisiae ORF YMR220w [
BRG8 : phosphomevalonate kinase ]"
/evidence=not_experimental
<278..>972
/note="similar to Saccharomyces cerevisiae ORF YBR166c [
TYR1 : prephenate dehydrogenase (NADP+)]"
1 putative frameshift(s)"
/evidence=not_experimental

ORIGIN
Query Match 5.5%; Score 71.8; DB 9; Length 973;
Best Local Similarity 74.0%; Pred. No. 1.9e-07;
Matches 91; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy      5 CAAAGCATTTAGTGCACCTGGAAAGCATTTCTGCTGATGATTTGTTCTTGAGC 64
Db      141 CTAGAGCATTCCTGCTCTCTGAAAAAGCATTTTGGCGGTTGATACCTCGTATAGAC 82
Qy      65 CAATTATGATGCTTATGTGACAGCATTTGTCATCAAGATGATGCAATTATACACCA 124
Db      81 CAATATACATCTTATGTGACAGCGCTTTCTTCAAGATGATGCAATTGTTAAAAATA 22
Qy      125 AAG 127
Db      21 GAG 19

RESULT 4
LOCUS   CENS06WQJ          917 bp    DNA          linear    GSS 06-JUL-2001
DEFINITION
T3 end of clone AX0AA034D03 of library AX0AA from strain CBS 7064
of Pichia farinosa, genomic survey sequence.
ACCESSION
AL418769
VERSION
AL418769.1 GI:12201829
KEYWORDS
GSS.
SOURCE
Pichia farinosa
ORGANISM
Pichia farinosa
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE
1 (bases 1 to 917)
Soulciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,

```

TITLE
Saurin, N., Tekala, F., Toffano-Nioche, C., Mesolowski-Louvel, M.,
Wincker, P. and Weissenbach, J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies

JOURNAL
MEDLINE
20584721
PUBMED
1152876
REFERENCE
2 (bases 1 to 917)
de Montigny, J., Spehner, C., Souciet, J., Tekala, F., Dujon, B.,
Wincker, P., Artiguenave, F. and Peller, S.
Genomic exploration of the hemiascomycetous yeasts: 15. *Pichia*
sorbitophila
FEMS Lett. 487 (1), 87-90 (2000)

JOURNAL
MEDLINE
20584725
PUBMED
1152890
REFERENCE
3 (bases 1 to 917)
Genoscope.
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: *Saccharomyces bayanus* var. *uvarum*, *Saccharomyces*
exiguus, *Saccharomyces servazii*, *Zygosaccharomyces rouxii*,
Saccharomyces kluyveri, *Kluyveromyces thermotolerans*, *Kluyveromyces*
angusta, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,
Candida tropicalis and *Yarrowia lipolytica*. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
the other extremity of this insert.

COMMENT

FEATURES
source
Location/Qualifiers
1..917
/organism="Pichia farinosa"
/mol_type="Genomic DNA"
/strain="CBS 7064"
/db_xref="taxon:4920"
/clone="AX0A034D03"
/clone_11b="AX0AA"
/note="end : T3"

ORIGIN

Query Match
Best Local Similarity 5.1%; Score 65.6; DB 9; Length 917;
Matches 133; Conservative 4; Mismatches 98; Indels 1; Gaps 1;
QY 720 TAACTGGGAATCAAAATGAAAGATGATCATTACCATAGGAAATCAATTATTATGGG 779
DB 94 TGAATCGCATTTCAAAATGCAAAATGCGCTTGGC-GCCASATCAAGCTGTGATGG 152
QY 780 TGAAGTCAAGGGTGGCTCAGAAACACCCAAATTTGTATTCAGAGTACTCCAATGAAAA 839
DB 153 ARAAGTTCAGGTGGTGAAGAGAACCCCAACTGTATCTGTAGTGTCAATAGGAA 212
QY 840 GGAAGGCGAAGAAAGCTCTGTGTGTATGACCACTTAATAGTGCATTTACGTT 899
DB 213 GCGAAGCCAGTCGAGGCGGAGAAAMTATTGAGATTTTAATAAGCCAAATGGTGGCT 272
QY 900 TATGAAGATTGAGGAAATGCTGTAATAATAGACTCAGACCCAGAGACTTATA 955
DB 273 CATTCATCATTTGTGCTACTTACAGAACTCCAGCGGAAACCCAGCCGCTACA 328

RESULT 5
CF824043/c
LOCUS
DEFINITION
CF824043 920 bp mRNA linear EST 01-APR-2004
EST701425 Coccidioides posadasii saprobic phase cDNA library, 2 to
4 kb Coccidioides posadasii cDNA clone CIDAP12 3' end, mRNA
sequence.
CF824043
VERSION
CF824043.1 GI:45930100
KEYWORDS
SOURCE
Coccidioides posadasii

ORGANISM

Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Ouygenales; Microsporitic Ouygenales; Coccidioides.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
20584721
PUBMED
1152876
COMMENT
1 (bases 1 to 917)
Gardner, M.J. and Cole, G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags
Unpublished (2003)
Other ESTs: EST701426
Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org

FEATURES
source
Location/Qualifiers
1..920
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDAP12"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_11b="Coccidioides posadasii saprobic phase cDNA
library, 2 to 4 kb"
/note="Vector: pXpress 1; Site 1: Not I; Site 2: Eco RV;
Coccidioides posadasii saprobic phase cDNA library, size
fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match
Best Local Similarity 4.9%; Score 63.8; DB 7; Length 920;
Matches 310; Conservative 0; Mismatches 312; Indels 39; Gaps 2;
QY 523 GATATTTTGACACAGCTTGACAGATTGCACTTTATAGCCAAAAAATGATGATCT 582
DB 897 GATTAACCTATTAACCTGCTCAAGCTCTCACTGTCGCCGCAAGGTAAGTGGCTCT 838
QY 583 GGGTTGATGTGCACTGCAATTTATAGTCTGATGATGATTAAGATTTACGCCACT 642
DB 837 GGAATTTGACATGGCGGCGCGGTGACGATCTTCTTAACTGCTGTTCTCCACATCT 778
QY 643 TTGATAAATGACGCTTCAAGTTCTGAAAGTATCTGAGAAGTCTCCAC----- 695
DB 777 GTTTCACGACCTGGGCGAGGTGGCTCCCGAGTTGAAACAGCGCCATTATACAGTT 718
QY 696 -----AGAGTGAATAAATGATTAAGTAAGTAACTGGGAATTCAAACATGAAGATGACA 750
DB 717 GTTGAAGACCTCAATACCAAAAAACCTGGGACACAGAGTGTGATTTGATTTAA 658
QY 751 TTACCATTCAGCAATCAAGTTATTAAAGGTGACGTCAAGGGTGTCAAAACACCCAA 810
DB 657 CTCCCAAGGGGCTCCAAATGTCTTTTGCAGCTGACGTGATCGGGTTCCGACCCCGCC 598
QY 811 TTGGTATTCAGAGTACTCCATATGAAAAAGAAAAAGCCAGAAAGAAAGCTCTGTGTAT 870
DB 597 ATGGTTAAAAAGGTGTGAGTGGCGGAAACGAAACCGGGGACGCAAGCTCTCTCG 538
QY 871 GACAGCTTAATAGTCCCAATTTACGTTATGAAGAAATGAGGAAATCGTGAATAA 930
DB 537 ACGGGGTACAGGGAACATGAATAAGCTTGCCTCGAGCTCAAGCAATGGCAGCAAAAC 478
QY 931 TACGATTCAGACCCAGAGACTTATTAAGAGTTAAGTTCATCTGTGAGCTTTGACT 990
DB 477 CGC-----AATGATGCTCAATTAATTAAGTGAACCTCAGC 445
QY 991 GTTGCATTAAGAAATGAGAAAGGTTTACAGCAATTAACAAAAATCAGAGTTTCA 1050
DB 444 AACTGATCACTGCACTGAGATGATTAAGTCAATGACGAAGAAATCGGGGTTCC 385
QY 1051 ATTGAACCTGATGTCAAACCAAGTTGTTGACCGTTGTCAAGAGATTCCTGTTGT 1110

Db 384 ATTGAACCTGAGTGCAGACAGATTATTGAATCCCTTTTCGAAAAATTGACGGGGTCATT 325
 QY 1111 GGTGCTGTGCTTCCAGTCTCTGTGATACATGATGATGATGATGATGATGATGATGAT 1170
 Db 324 GGTGCGCTTGTTCCTCCGCGCGCGGCGATATGATGATGATGATGATGATGATGATGAT 265
 QY 1171 C 1171
 Db 264 C 264

RESULT 6
 C0014330/c 846 bp mRNA linear EST 09-JUN-2004
 LOCUS EST802665 Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
 DEFINITION kb Coccidioides posadasii cDNA clone CIEC383 3' end, mRNA sequence.
 ACCESSION C0014330.1 GI:48521219
 VERSION C0014330.1 GI:48521219
 KEYWORDS EST.
 SOURCE Coccidioides posadasii
 ORGANISM Coccidioides posadasii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Onygenales; Mitosporic Onygenales; Coccidioides.
 1 (bases 1 to 846)
 GARDNER/M.J. and COLE/G.T.
 ANALYSIS of gene expression in Coccidioides posadasii mycelia and
 spherules via expressed sequence tags
 UNPUBLISHED (2003)
 JOURNAL Other ESTs: EST802666
 COMMENT Contact: Gardner MJ
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 3519
 Fax: 301 838 0208
 Email: gardner@tigr.org.

FEATURES
 source location/Qualifiers
 1..846
 /organism="Coccidioides posadasii"
 /mol_type="mRNA"
 /strain="C735"
 /db_xref="taxon:199306"
 /clone="CIEC383"
 /dev_stage="spherules"
 /lab_host="E. coli DH10B, T1 phage resistant"
 /note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
 Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
 kb"

ORIGIN
 Query Match 4.8%; Score 62.2; DB 7; Length 846;
 Best Local Similarity 46.7%; Pred. No. 6e-05;
 Matches 309; Conservative 0; Mismatches 313; Indels 39; Gaps 2;

QY 523 GATATTTGACACAGCTGACAGATTGACATTTGATGCCCCAAAAAGATAGAGATCT 582
 Db 821 GATTAACATCAATACCTGCTCAAGCTGCTCACTGCGCCCAAGATAAGTAAAGCTCTT 762
 QY 583 GGGTTGATGTTGCAACATGCAATTTATGCTGATGATGATGATGATGATGATGATGATGAT 642
 Db 761 GGAATTTGACATTCGCGCAGCCGTCGATCTTGTCTTTACCGTCTTTCACCATCT 702
 QY 643 TTGATTAATGACGTGTTTCAAGTTCAGAAAGATGATCCCTGAGAGATTCCCGAC----- 695
 Db 701 GTTCTCAGGACCTGGCGAGGTGGCTCCCGCACTTTTGAACAGCGCTATTATTAAGTT 642
 QY 696 -----AGAGTTAAAAAATTGATTTGAAGTAATCTGGGAATTCAAACATGAAGATGTACA 750
 Db 641 GTTGAAGACCTCAATACCAAAAACTTGGACACAGAGTGTGTAGACTTTGATTTAA 582
 QY 751 TTACCAATACGAATCAAGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 810

Db 581 CTCACAGGGGGCTCCAAATGCTCTTTGCGACGTGAGTGGCTTCCAGACCCCGGC 522
 QY 811 TTGATATCAAGATCACTCCATGATAAAAAAGCAAGCAAGAAAGCTCTGTGTAT 870
 Db 521 ATGTTAAAAAGTCTGTGAGTGTGGGGAACGAACCCGGGAGCGCAGACTGCTCG 462
 QY 871 GACCACTTAATATGTCGCAATTTACATTTATGAAAGATGAGGAAATGCGTAAGAA 930
 Db 461 ACGGGGTACAGCGGAACATGAAAAGCTTCGCTGAGCTCAAGCGATTGGACACAAC 402
 QY 931 TACGACTAGACCCAGACATTAATTAAGAGTTAGATCTCTGTGAGCTTTGACT 990
 Db 401 CGCAAC-----GTAATCCAGAACTATGATGACTCAGC 369
 QY 991 GTTGGATTAAGACATGAGAAAGGTTTACAGGATTAACAAAAATCAGAGTTCCA 1050
 Db 368 AACTGATCACTCGCATCTAGATGTGGAATTAGTCAATGACAGAGAAATCGGGGTTCC 309
 QY 1051 ATTGAACCTGATGTCGAACCCAGTTGTGACCGTTGTCAAGAGATTCTGTGTGTT 1110
 Db 308 ATTGAACCTGAGTGCAGACAGAAATTATGAAATGCTTTTCGAAATGACGGGTCATT 249
 QY 1111 GGTGCTGTGCTTCCAGTCTGTGATACATGATGATGATGATGATGATGATGATGATGAT 1170
 Db 248 GGTGCGCTTGTTCCTCCGCGCGCGGCGATATGATGATGATGATGATGATGATGATGAT 189
 QY 1171 C 1171
 Db 188 C 188

RESULT 7
 C0008271/c 922 bp mRNA linear EST 09-JUN-2004
 LOCUS C0008271 Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
 DEFINITION kb Coccidioides posadasii cDNA clone CIEB521 3' end, mRNA sequence.
 ACCESSION C0008271
 VERSION C0008271.1 GI:48515160
 KEYWORDS EST.
 SOURCE Coccidioides posadasii
 ORGANISM Coccidioides posadasii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Onygenales; Mitosporic Onygenales; Coccidioides.
 1 (bases 1 to 922)
 GARDNER/M.J. and COLE/G.T.
 ANALYSIS of gene expression in Coccidioides posadasii mycelia and
 spherules via expressed sequence tags
 UNPUBLISHED (2003)
 JOURNAL Other ESTs: EST796607
 COMMENT Contact: Gardner MJ
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 3519
 Fax: 301 838 0208
 Email: gardner@tigr.org.

FEATURES
 source location/Qualifiers
 1..922
 /organism="Coccidioides posadasii"
 /mol_type="mRNA"
 /strain="C735"
 /db_xref="taxon:199306"
 /clone="CIEB521"
 /dev_stage="spherules"
 /lab_host="E. coli DH10B, T1 phage resistant"
 /clone_lib="Coccidioides posadasii spherule cDNA library,
 0.4 to 2.3 kb"
 /note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
 Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
 kb"

ORIGIN
 Query Match 4.8%; Score 62.2; DB 7; Length 922;
 Best Local Similarity 46.7%; Pred. No. 6.1e-05;

Matches 309; Conservative 0; Mismatches 313; Indels 39; Gaps 2;

QY 523 GATATTTTGCACAACTTGCACAGATTCAGATTGTTATGCCCAAAAAAGATGAGATCT 582
 DB 836 GATTAACATCAATACCTGCTCAGCTCTCAGCTGCGCGCAAGTAAGTAGGCTCT 777
 QY 583 GGGTTGATGTTGCACTGCACATTTATGCTGATGTTGATTAAGAAATTCAGCCAGCT 642
 DB 776 GGAATTTGACATTTGCGCGACCGCTGACGATCTTTCCTTAACGTTGCTTCCATCT 717
 QY 643 TTGATTAATGACGTTTTCAGTTCTGAAAGATATCTTGAAGATTTCCAC----- 695
 DB 716 GTTCTCAGCGACCTGGGAGGTTGGCTCCCCGAGTTTGAACAGCGCCATTATACAGTT 657
 QY 696 -----AGAGTTGAAAAAATTGATTGAAGTAACTGAGAAATTCAAACATGAAGATGTACA 750
 DB 656 GTTGAAGCTCAATATACGAAAAAATCTTGGGACACAGATGTGTACATTTGATTTAA 597
 QY 751 TTACCATACGGAATCAATTATTAATGGTGACGTCAAGGTTGCTCAGAAAAACCCCAA 810
 DB 596 CTCCCAAGGGGCTCCAAATGCTCTTGGAGCTGACCTGCGGTTCCAGACCCCGGC 537
 QY 811 TTGGTATCAGAGTATCTCCATGTGAAAAAGAAAAAGCCAGAAAGCTCTGTTGTAT 870
 DB 536 ATGGTTAAAAAGGTGTGAGTGCGCGGAAACGAACCCGGGACCGACGCTCTCTGG 477
 QY 871 GACCACTTAATAGTCCCAATTGACATTTAGAAAGATTAAGGAAATGCGTGAATA 930
 DB 476 ACCGGGCTACAGGGACAAATGAAAGCTTGCGCTCAGCTCAAGCAATTTGGACAAAC 417
 QY 931 TACGACTCAGACCCAGACCTTAATTAAGAGTTAGATCATCTGTTGAGCCTTGACT 990
 DB 416 CGC-----AATGATGCAAGATTTATGTAACCTCAG 384
 QY 991 GTTGGATTAAGACATCAGAAAAAGGTTTACAGCATTTACAAAAATCAGAGTTCCA 1050
 DB 383 AACTTGATCATCTCGCAGTACAGATGTGATTAAGTCAATGACGAAGAAATCGCGGTTCC 324
 QY 1051 ATTGAACCTGATGTCGAACCCAGTTGTTGACCGTTGTCAAGAGATTCCTGTTGTGT 1110
 DB 323 ATTGAACCTGAGTGACAGACAGATTAATGATGCTTTTCCAAATTAACGCGGTCATT 264
 QY 1111 GGTGGTGTGTTCCAGGTGCTGAGATTAAGATGCAATGCTGTATTAAGTTGGAAT 1170
 DB 263 GGTGGCGTGTGTTCCGCGCGCGGCGGATGATGACCTGCTTCTGTTAAAGATGAT 204
 QY 1171 C 1171
 DB 203 C 203

RESULT 8
 LOCUS AZ931444 488 bp DNA linear GSS 01-Apr-2001
 DEFINITION 474.dhz84b12.s1 Saccharomyces unisporus NRRL Y-1556 Saccharomyces
 accession genomic clone 474.dhz84b12.s1, genomic survey sequence.
 VERSION AZ931444
 KEYWORDS GSS.
 SOURCE Saccharomyces unisporus
 ORGANISM Saccharomyces unisporus
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 REFERENCE 1 (bases 1 to 488)
 AUTHORS Gish, W.R., Hillier, L.W., Fulton, L., Graves, T., Miner, T.,
 Surveying Saccharomyces genomes to identify functional elements by
 comparative DNA sequence analysis
 JOURNAL Unpublished (2001)
 COMMENT Contact: Johnston M
 Department of Genetics
 Washington University Medical School
 Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA

FEATURES
 source
 1..488
 /organism="Saccharomyces unisporus"
 /mol_type="genomic DNA"
 /strain="NRRL Y-1556 (CBS 398)"
 /db_xref="taxon:27294"
 /clone="474.dhz84b12.s1"
 /clone_lib="Saccharomyces unisporus NRRL Y-1556"
 /note="Random genomic sequence"

ORIGIN

Query Match 4.7%; Score 60.6; DB 8; Length 488;
 Best Local Similarity 54.7%; Pred. No. 0.00013;
 Matches 168; Conservative 0; Mismatches 129; Indels 10; Gaps 2;

QY 991 GTTGGATTAAGACATCAGAAAAAGGTTTCAAGCATTTACCAAAAAATCAGAGTTCCA 1050
 DB 5 GCTCCATTTACAAATTTACAGAAATTAATTTAGAACATTCACAAAAAGATCGGTGAC 64
 QY 1051 AT-TGAACCTGATGTCGAACCCAGTTGTTGACCGTTGCAAGATTCCTGTTGTGT 1109
 DB 65 ATGTGGAACATGTGTACAAACAAATTCCTAAACATTTGTCATTAATGTTGTCAT 124
 QY 1110 TGTGTGTGTGTTCCAGGTCCTGTGTGATGATGATGATGATGATGATGATGATGAT 1169
 DB 125 AGGTGTGTGTGTTCTCGTGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 184
 QY 1170 TCAAGTGGGAATTTTAAACAGAAAACTTTGAAATTCAGATTAATTTCTAATGTTTA 1229
 DB 185 -----AATGATTAATTAAGATGATTAAGAAACAAACATTTTCAAAAAATGTCAC 235
 QY 1230 CTGGGTTGATTTGGAAGGAAACGAAGGTGTCTTGAAGAAAAACAGAGATTTAT 1289
 DB 236 TTGATGAGACGTTTAAACAAACAAATTTGATTAAGAAAGAAAAATCBAACATTTATGA 295
 QY 1290 AGGTTTA 1296
 DB 296 AGATCTTA 302

RESULT 9
 LOCUS BZ304881/c 482 bp DNA linear GSS 31-OCT-2002
 DEFINITION KD3193.p1 Kluyveromyces delphensis Random Genomic Library
 sequence.
 VERSION BZ304881
 KEYWORDS GSS.
 SOURCE Kluyveromyces delphensis
 ORGANISM Kluyveromyces delphensis
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 REFERENCE 1 (bases 1 to 482)
 AUTHORS Wong, S., Fares, M.A., Zimmermann, W., Butler, G., and Wolfe, K.H.
 Evidence from comparative genomics for a complete sexual cycle in
 the 'asexual' pathogenic yeast *Candida glabrata*
 JOURNAL Genome Biol. 4 (2), K10 (2003)
 MEDLINE 22508158
 PUBMED 12620120
 COMMENT Contact: Wong S
 Department of Genetics, Smurfit Institute
 Trinity College Dublin
 Dublin 2, Ireland
 Tel: 353 1 6082319
 Fax: 353 1 6798558
 Email: swong@tcd.ie
 Class: plasmid ends.
 Location/Qualifiers

FEATURES

Db	158	AATGAGAAATGGGTAAAATTGCATGTACCATTTGAACAATGAACAACTGAATT	217
Oy	1077	GTTGACCCTTGTTCAGAAGATTCTCGTTGTGTTGGTGGTGTGTTCCAGTGTGGTG	1136
Db	218	GGCGATCAATCAATGTCAATCAATGGTTGTATTTGCGTGGGTACTCGTGTGGTG	277
Oy	1137	ATAAGATGCAATAGCTGTATTAGTGTGGAAATCAAGTGGGAAATTTTAA	1187
Db	278	TTTCGATGCTTTATTGCTGTATCGTTATCATTCAATCAAGATGAATTTAAAAA	328
RESULT 11			
AU268448		479 bp	mRNA linear EST 26-APR-2004
LOCUS			
DEFINITION	AU268448 VS Dictyostelium discoideum cDNA clone VS1284 5', mRNA sequence.		
ACCESSION	AU268448		
VERSION	AU268448.1		
KEYWORDS	GI:20527246		
SOURCE	EST.		
ORGANISM	Dictyostelium discoideum		
REFERENCE	Dictyostelium discoideum		
AUTHORS	Dictyostelium discoideum Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium. 1 (bases 1 to 479) Urushihara,H., Morio,T., Saito,T., Kohara,Y., Koriki,E., Ochiai,H., Meda,M., Williams,J.G., Takuchi,I. and Tanaka,Y. Analyses of cDNAs from growth and slug stages of Dictyostelium		
TITLE	dictyostelium		
JOURNAL	Nucleic Acids Res. 32 (5), 1647-1653 (2004)		
COMMENT	Contact: Hideko Urushihara Institute of Biological Sciences University of Tsukuba 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan Tel.: 81-298-53-4664 Fax: 81-298-53-6614 Email: hideko@iol.tsukuba.ac.jp.		
FEATURES	location/Qualifiers		
SOURCE	1..479		
	/organism="Dictyostelium discoideum"		
	/mol_type="mRNA"		
	/strain="AX4"		
	/db_xref="taxon:44689"		
	/clone="VS1284"		
	/sex="mat A"		
	/dev_stage="vegetative"		
	/clone_1fb="VS"		
ORIGIN			
Query Match	4.2%; Score 55; DB 1; Length 479;		
Best Local Similarity	52.4%; Pred. No. 0.0036;		
Matches 121; Conservative	0; Mismatches 110; Indels 0; Gaps 0;		
Oy	957	TAAAGATTGATCATCTGTGAGCCTTTGACCTGTTCGATTAAAGACATCAGAAAAGG	1016
Db	86	TAAATAATTTGTAATACTTTAATTAACAATTAGAGATCATTTTTAAAGATPAATCATTT	145
Oy	1017	GTTACAAGCATTAACACAAAATCAGAGTTTCCAATTAACCTGATGCCAACCCAGTT	1076
Db	146	AATGAGAAATGGGTAAAATTGCTGATGTATCACTTGAACCAATGAMCAAACCTGATT	205
Oy	1077	GTTGACCGTTGTCAAGATTTCTCGTTTGTGTTGGTGGTGTGTTCCAGTGTGGTG	1136
Db	206	GGCTGATCAATCAATGTCAATCAATGGTTGTATTTGCTGCTGTACTCGTGTGGTG	265
Oy	1137	ATAAGATGCAATAGCTGTATTAGTGTGGAAATCAAGTGGGAAATTTTAA	1187
Db	266	TTTCGATGCTTTATTGCTGTATCGTTATCATTCAAGATGAATTTAAAAA	316
RESULT 12			
BJ343297		717 bp	mRNA linear EST 07-MAR-2002
LOCUS			
DEFINITION	BJ343297 Dictyostelium discoideum cDNA library, Af Dictyostelium discoideum cDNA clone dda22j18 3', mRNA sequence.		

```

FEATURES
source
1. 1101
location/Qualifiers
1. 1101
location/badpath: med. buf.fallo.edu/drosophila_bac.htm.
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_1lb="RPCT-98"
/note="end : IET3"

```

Query Match	4.1%;	Score 53.4;	DB 9;	Length 1101,
Best Local Similarity	15.0%;	Pred. No. 0.013;		
Matches	85;	Conservative	258.	Mismatch-

[illegible]

RESULT 14	
AU052799	
LOCUS	
DEFINITION	
AU052799	322 bp mRNA
AU052799 Dictyostelium discoideum SL (H.Urushihara)	linear EST 28-APR-1995
discoideum cDNA clone SLF121, mRNA sequence.	Dictyostelium

ACCESSION AU052799.1 GI:4701282
VERSION AU052799.1
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 322)
AUTHORS Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshida, R., Mitsu, B. N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H., Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
TITLE Developmental cDNA in Dictyostelium discoideum
JOURNAL Unpublished (1998)
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp
PROJECT = Dictyostelium discoideum cDNA project in Japan.

FEATURES
source
1..322
location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLF121"
/dev_stage="slug"
/clone_11b="Dictyostelium discoideum SL (H. Urushihara)"

ORIGIN
Query Match 4.0%; Score 51.6; DB 1; Length 322;
Best Local Similarity 58.4%; Pred. No. 0.026;
Matches 90; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1034 AAAATCAGAGGTTCCATTTGAACCTGATGCCAACCCAGTTGTGACCGTTGTCAAG 1093
DB 4 AAATGCTGATGATCATTTGGAACCAATGAAACCAATGATGCTGATCATCAATGT 63
QY 1094 AGATTCCTGT 1153
DB 64 CAATCAATGT 123
QY 1154 TATAGTGTGTGAAATCAAGTGGGAAATTTAA 1187
DB 124 CTATCTTATCTCAATCAAGATGAATTAATAA 157

RESULT 15
LOCUS CO142555 360 bp mRNA linear EST 17-JUN-2004
DEFINITION EST137226 Aspergillus flavus Normalized cDNA Expression Library
ACCESSION CO142555
VERSION CO142555.1
KEYWORDS EST.
SOURCE Aspergillus flavus
ORGANISM Aspergillus flavus
REFERENCE 1 (bases 1 to 360)
AUTHORS Eukariotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eutectiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
TITLE Yu, J., Whitelaw, C.A., Niernan, W.C., Bhatnagar, D. and Cleveland, T.E.
Aspergillus flavus expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops
JOURNAL FEMS Microbiol. Lett. (2004) In press
COMMENT Contact: Yu J
Food and Feed Safety Research Unit
USDA/ARS, Southern Regional Research Center
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
Tel: 504 286 4405
Fax: 504 286 4419
Email: jinyu@rrc.ars.usda.gov
Contact Dr. Yu at USDA/ARS SRRC (jinyu@rrc.ars.usda.gov) for clone

Information
PCR Primers
FORWARD: M13F
BACKWARD: M13R
Seq primer: M13 Forward
POLYA=No.

FEATURES
source
1..360
location/Qualifiers
/organism="Aspergillus flavus"
/mol_type="mRNA"
/strain="NRRL 3357"
/db_xref="taxon:5059"
/clone="NAF095"
/sex="asexual mycelia"
/cell_type="mycelia"
/dev_stage="developmental stages from 18 to 96 hours"
/lab_host="B. coli DH10B T1 resistant cells"
/clone_11b="Aspergillus flavus Normalized cDNA Expression Library"
/note="Vector: pBluescript (SK+) (Stratagene), antibiotic selection marker: Carbenicillin; Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

ORIGIN
Query Match 3.8%; Score 49.2; DB 7; Length 360;
Best Local Similarity 58.0%; Pred. No. 0.12;
Matches 87; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1028 TAACCAAAATCAAGGTTCCATTTGAACCTGATGCCAACCCAGTTGTGACCGTT 1087
DB 56 TGACCCGCAAGCGGAGATGCCATTTGAGCCCGGGTACAGACAGAACTGCTTGATGCCG 115
QY 1088 GTCAAGAGATTCCTGT 1147
DB 116 TGTCAAGCCGTTACGAGATGATTTGT 175
QY 1148 TAGCTGTATTAGTGTGGAATCAAGTGG 1177
DB 176 TTGCTGTCTGTATCCGCGATGACAGAGG 205

RESULT 16
LOCUS CO134615 412 bp mRNA linear EST 17-JUN-2004
DEFINITION EST829286 Aspergillus flavus Normalized cDNA Expression Library
ACCESSION CO134615
VERSION CO134615.1
KEYWORDS EST.
SOURCE Aspergillus flavus
ORGANISM Aspergillus flavus
REFERENCE 1 (bases 1 to 412)
AUTHORS Eukariotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eutectiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
TITLE Yu, J., Whitelaw, C.A., Niernan, W.C., Bhatnagar, D. and Cleveland, T.E.
Aspergillus flavus expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops
JOURNAL FEMS Microbiol. Lett. (2004) In press
COMMENT Contact: Yu J
Food and Feed Safety Research Unit
USDA/ARS, Southern Regional Research Center
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
Tel: 504 286 4405
Fax: 504 286 4419
Email: jinyu@rrc.ars.usda.gov
Contact Dr. Yu at USDA/ARS SRRC (jinyu@rrc.ars.usda.gov) for clone
PCR Primers


```

/clone="NAGF24"
/sex="asexual mycelia"
/cell_type="mycelia"
/dev_stage="developmental stages from 18 to 96 hours"
/lib_host="S. coli DH10B T1 resistant cells"
/clone_lib="Aspergillus flavus Normalized cDNA Expression
Library"
/note="Vector: pBluescript (SK+) (Stratagene), antibiotic
selection marker: Cambericillin; Site_1: NotI, at the 5
prime end; Site_2: EcoRI, at the 3 prime end; This
normalized cDNA expression library was constructed using a
mixture of mycelial cells grown under eight different
medium conditions and harvested at 5 time points (18, 24,
48, 72, 96 hours). The poly-A sequence was trimmed off
before ligating to vector."

```

before ligating to vector."

ORIGIN

Query Match

Query Match	3.7%	Score 47.8;	DB 5;	Length 775;
Best Local Similarity	60.3%;	Pred. No. 0.33;		
Matches	79;	Conservative	0;	Mismatches

[illegible][illegible]

203 TTGCTGGTGGATCCGGGATGACCAAGGAGG 292

RESULT 21
C0152593

DEFINITION

EST827646	Aspergillus flavus	447 bp	MRNA	linear	EST 17-JUN-2004

ACCESSION
VERSION

CO152593 5' end, mRNA sequence.

**RECORDS
SOURCE**

EST.
Asperail Inc f1 3/20/2000

ORGANISM

Aspergillus flavus

REFERENCE

1 (bases 1 to 447)

AUTHORS

Yu, J., Whitelaw, C.A., Nierman, W.C. Rhinoceros

JOURNAL.

genes with putative roles in aflatoxin contamination of FEMS Microbiol.

seq primer: M13 Forward
POLYA=No.

```

source
location/Qualifiers
1..447
/organism="Aspergillus flavus"
/mol_type="mRNA"
/strain="NRRL 3357"
/db_xref="taxon:5059"

```

FEATURES

Source

```

/organism="Aspeggillus flavus"
/mol_type="mRNA"
/strain="NRRL 3357"
/db_xref="taxon:5059"
/clone="NAFCF88"
/sex="asexual mycelia"

```


[illegible]

ANIMORPHS	SOURCE
ORGANISM	EST.
	<i>Coccidioides posadasii</i>
	<i>Coccidioides posadasii</i>
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes
	Onygenales; mitosporic Onygenales; Coccidioides
REFERENCE	1 (bases 1 to 803)
ANIMORPHS	

Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@igb.org

```

FEATURES
    source
        location/Qualifiers
            1..803
                /organism="Coccidioides posadasii"
                /mol_type="mRNA"
                /strain="C735"
                /db_xref="taxon:199306"
                /clone="CIEC383"
                /dev_stage="spherules"
                /lab_host="E. coli DH10B, T1 phage resistant"
                /clone_1lb="Coccidioides posadasii spherule cdna library
0.4 to 2.3 kb"
                /note="Vector: pExpress 1; Site 1: Not I; Site 2: Eco RV
Coccidioides posadasii spherule cdna library, 0.4 to 2.3
kb"

```

Query Match 3.7%; Score 47.6; DB 7; Length 803;
Best Local Similarity 52.4%; Pred. No. 0.38;
Matches 132; Conservative 0; Mismatches 114; Indels 6; Gaps 1

RESULT 26	
CNS0106x	
LOCUS	
DEFINITION	CNS0106x 1101 bp DNA linear GSS 26-UTL-1999
ACCESSION	BACN03K20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION	AL098595
KEYWORDS	AL098595.1 GI:5610206 GSS.
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster

TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
Determination of this RAC and comment

collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk/>. This *Drosophila melanogaster* BAC library (Dros BAC) was made by Alain Billaud at CSH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelosAC11.

```

FEATURES
SOURCE
    Location/Qualifiers
    1..1101
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="tsg:17227"
    /clone="BACN03K20"
    /clone_1kb="DrosBAC"
    /plasmid_id="pbeloBAC11"
    /note="end : T7"
ORIGIN

```

Query Match	3.7%;	Score 47.6;	DB 9;	Length 1101;
Best Local Similarity	18.9%;	Pred. No. 0.41;		
Matches	96;	Conservative 198.	Mismatches 0.	

		Accession	Size	Insets	1; Gaps	1
OY	514	ACGAATPAAGATATTTCGACAAGCTTGCAACAGTTGCACATTTGTATGCCAAAAAAG	573			
Db	555	AGGRITMGKVKBKHHHMTTGGTKKKTCTCHKTYYTKWTWIMTKAGSTKXTGTPDK	614			
OY	574	ATAGAATCGGGTTGATGTTGSCATGCAATTATGCTGTATGTAATATGAAGATTT	633			
Db	615	TGATMAATGTAATATGTAATKAATKRADTDGTATATATAAKKKARXNDXTAADAAKW	674			
OY	634	CAGCCAGCTTTGATPAAATGACGTTTTCAGGTTCTGAAGAAAGTACTCGAAGATTTCCC	693			

Db	675	DKATKAKKKA	KKKAAKAKAA	AKTATG	KARAKAKAKAKA	KTAKARAK	MDTATWTAT	TAAD	734				
Qy	694	ACAGATTG	AAAAAATTG	ATTGAA	AGTAAGT	AACTGG	GAATTCA	AACATGA	AAAGATG	TACATTA	753		
Db	735	AAADKGA	KDKAKAKA	KADAPDK	RNRDAD	KDKDKR	KAKAKAA	GAADAD	DADADAKA	AAAKA	794		
Qy	754	CCATACGA	AAATCA	ATTATTA	TATG	GGTG	CGCTCA	AGGGT	GGCTC	GAAACA	CCCAATTG	813	
Db	795	KADAAAD	DDRG	DGGDKK	KRAKDK	RRKKKK	KKDK	KAMGDK	KAKKDK	KA	AKAKADA	AGAPAKAR	854
Qy	814	GTATCA	CGAGTA	CTTCCA	TGAAAA	AGAAAA	CCCA	GAAAGCT	CTGTGT	GTATG	AC	873	
Db	855	RAGKDK	KAKADAKA	KAKAKAKA	KDDDAK	ATAKAT	AKATKAT	KAKDKAKA	KKKKKK	KDKADAKA	KA	914	
Qy	874	CAGCTT	ATAGTGC	CAATTT	TACAGT	TATG	-A	AGGAATT	GAGGGA	ATGCGT	GAAAAATA	932	
Db	915	KAKAKD	KDDDDDK	DKAD	KKKKKK	KORAK	KDKDDDK	KDKAD	KDKDDDK	ADAK	974		
Qy	933	CGACTCA	GACCCAG	AGACTT	ATATTA	AAAGAT	TGATCAT	TTCTGT	GTGAGC	CTTGAC	CTGT	992	
Db	975	AWAARAK	DADADAKA	KAKAKAKA	KADADDA	ADAKA	KADAKA	KA	KAKADDA	AAKAKAKA	KADAVAA	1034	
Qy	993	TGCGATT	AAAGAC	ATCAG	AAAAAG	GGTTA	1020						
Db	1035	DADAKAKA	KRAKAKAD	ADAAKAKA	KA	KA	KA	KA	KA	KA	KA	1062	

RESULT 27	CNS0182P	LOCUS	DEFINITION
	CNS0182P	1101 bp	DNA linear GSS 26-JUL-1999
			Drosophila melanogaster genome survey sequence SP6 end of BAC
			BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit
			fly), genomic survey sequence.

ACCESSION	AL108811
VERSION	AL108811.1
KEYWORDS	GI:5629115
SOURCE	GSS.
ORGANISM	<i>Drosophila melanogaster</i> (fruit fly)
	<i>Drosophila melanogaster</i>

REFERENCE
1 (bases 1 to 1101)

Genoscope - Centre National de Sequencage ::
Submitted (23-JUL-1999) Direct Submission
Genoscope.
AUTHORS
TITLE
JOURNAL

COMMENT

Determination of this BAC-end sequence was carried out as part of a

- Web : www.genoscope.cns.fr

BP 191 91006-ÉVRY cedex - FRANCE (E-mail : sege@genoscope.cns.fr)

collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES	source	location/Qualifiers
1.	.1101	
/organism="Drosophila melanogaster"		
/mol_type="Genomic DNA"		
/db_xref="taxon:7227"		
/clone="BACN37D10"		
/clone_11b="DrosBAC"		
/plasmid="pBelOBAC11"		
/note="end : SP6"		

ORIGIN	Query Match	3.6%	Score 46.6	DB 9	length 1101
	Best Local Similarity	20.9%	Pred. No. 0.75		
	Matches 99	Conservative 162	Mismatches 212	Indels 0	Gaps 0
QY	694	ACAGAGTGAATAAATTGTAAGTAACTCGGAATTCGAACATGAAGAATGTACATTA	753		

Db 576 AAAAADPTKORWDARGAAARAAAAAADAARFAAKTVAIAAAAAAAAAAAAAAAAAAAAA 635

Qy 754 CCATACCGAATCAAGTTATTATTAAGGTGACGCTCAAGGTGGCTCAAGAAACCCAAATTG 813

Db 636 GTTATTAAAAAAGGCMKTGKRAADODGGAAAAADWRRGRGRRBARBARAAAAAADAK 695

Qy 814 GTATCAGAGTACTCCAAATGAAAAAGAAAAAGCCAGAGAAAGCTCTGTTGTATGAC 873

Db 696 RAAAAAIAAAAAAAAAAADAAGRRKKKGGKGGKATKTAIAWAKGKDWGTATAWTW 755

Qy 874 CAGCTTATAGTGCATTTTCAAGTTATGAAGAAATTGAGGAAATTCGTGAIAAATAC 933

Db 756 DATWADTWKATTDPAKRAAAGRRKRAKTAARCGRRARRRRAWAAAGRRARAGARA 815

Qy 934 GACTCAGACCCAGAGACTTATATTAAAGTTAATATCTTGTGAGCCTTGACGTG 993

Db 816 RAARAADDDRWDMAAAAAAAAAAAAMTTWRBMDWDMDITRWDITTAAMWDDRA 875

Qy 994 GCGATTAAAGACATCAAGAAAAAGGTTACAAGCATTAACAAAAATCAGAGTTCAATT 1053

Db 876 RARRRRRRRRRRARRARRAADDTTKORMAATDTDKTTKTWTTDDDDWDRAKDRW 935

Qy 1054 GAACCTGATGCAAAACCCAGTTGTGTGAACGGTGTCAAGAGATTCCGTGTGATTTGGT 1113

Db 936 WAAARADGAWKRDRBARDDMAATAODDGDWKKWGGRGKGGKDKRWMDKGTGKKDDDDW 995

Qy 1114 GGTGTGGTTCAGAGTGTGCTGGAATACAGATCAATTAAGTTATAGTTTGA 1166

Db 996 DKTTTRDMMWTTTKMDWMDDBGRGKRWTRKRWGAMWRADMAARDDTDSKD 1048

RESULT 28	BO866497	716 bp	mRNA	linear	EST 14-Aug-2002
LOCUS	OG58C07.Y9.ab1	OG_ABCDI	lettuce	salinas	Lactuca sativa cDNA clone
DEFINITION	OG58C07, mRNA	sequence.			

ACCESSION	BO866497
VERSION	BO866497.1
KEYWORDS	GI:22251962
SOURCE	EST.
ORGANISM	<i>Lactuca sativa</i>
	<i>Lactuca sativa</i>

REFERENCES
1 (bases 1 to 716)

AUTHORS	TITLE
Kozik, A., Michelmore, R. W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolman, J., Stabaugh, M. S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.	Lettuce and Sunflower ESTs from the Compositae Genome Project

JOURNAL COMMENT
http://compgenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W. Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel.: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: ahorzi@atgsc.org [michelmore@vegmail.ucdavis.edu]
belongs to contlig_QC_Ca_contlig6356, see <http://cspdb.ucdavis.edu/>
for details.
row: C column: 07.
Location/Qualifiers
1..716

```

/organism="Lactuca sativa"
/moi_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QGC8C07"
/lab_host="E.coli"
/clone_id="QG-ABCDI lettuce salinas"
/vector="parCDNASTAB; The library was constructed

```

Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, transformationally cloned into a custom medium-copy vector and size biased. Details of each source of RNA and library construction can be obtained at <http://csgdb.ucdavis.edu/TAG-TISSE=Chemical> induction

```
440_3332=161A6CCEGG
```

Query Match	3.6%	Score	46.2;	DB	5;	Length	745;
Best Local Similarity	59.5%	Pred	No	0	05		

Score	Pred
59.58	

QY 51 AATTAAGATNTTTGGACAACGTTGCACAGATTGCACATTGTATATGCGCAAAAAAGATA 576
Db 562 AATCTTGATGTGGTGCATATTATATACCCAAATGGACACGTGTATGCTCAAGAAAAGTT
QY 577 GGAATCGGGTTGATGTGGCACTGCACATTATATGTGTGATATATGAAGATTTCAG 621
Db 622 GGGAGTGGAGTTGAATGTAGTTCAAGCTTTATGGAAGTCATCGTTATGTGAAGATTTCGA 636
QY 637 CCAAGCTTTGAT 647
Db 682 CCAACAAGTTAT 692

QY	517	AATPAAAGATATTTTGCACAAAGTTGCACAGATTTCACATTTGATATAGCCCAAAAAAAGATA	576
Db	562	AATCTTGATGTGTGTGCAATATATATACCCAAATGGCACACATGTAATGCTCAAGAAAAGTT	621
QY	577	GGATCTGGGTTTGATGTGTGCACATGCATTTATATGTGATGTATATATGAAGAATTTGAG	636
Db	622	GGAGAGGATTTGATGTATGATTGACCTGTTTATGGAAGTCATCGTATGTAAGATTTC	681
QY	637	CCAGGCTTTGAT	647
Db	682	CCACAAGTTAT	692

STAT 692

RESULT	29
BO870439	
LOCUS	
DEFINITION	BO870439 745 bp mRNA
ACCSSION	Q059423.yg.ab1 QG ABCDI lettuce salina Lactuca sativa
VERSION	Q059423.1 mRNA sequence.
KEYWORDS	BO870439 BO870439.1 GI:22256848
SOURCE	EST.
ORGANISM	Lactuca sativa
	Lactuca sativa
	Lactuca sativa
	linear EST 15-AUG-2007
	Lactuca sativa cDNA clone

RESULT 30
CNS017KC/c
LOCUS
DEFINITION
CNS017KC 1101 bp DNA linear GSS 26-UTL-1999
Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN37F10 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL108171
ACCESSION
AL108171.1 GI:5628475
VERSION
GSS.

01:306284/5

eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids, Asterales; Asteraceae; Cichorioideae; Cichoriaceae; Lactuca.
 1 (bases 1 to 745).
 Kozik, A., Michelmore, R. W., Kravetz, S. E., and Nelson, R. E. 2002.

SOURCE ORGANISM	REFERENCE
<i>Drosophila melanogaster</i> (fruit fly)	
<i>Drosophila melanogaster</i>	
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidae; Drosophilidae; Drosophila.	
1 (bases 1 to 1101)	

11071) *Drosophilidae*; *D.*

TITLE **JOURNAL** **COMMENT**

Jun, H., van Damme, M., Lavelle, D., Petrucci, M., Rieseberg, L.,
Ellison, P., Kolkman, J., Slabaugh, M. S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
lettuce and Sunflower ESTs from the Composite Genome Project
<http://compenomics.ucdavis.edu/>
Unpublished (2002)
Contact: Alexander Zaitchik
Unpublished (2002)

AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequençage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the

on of this BAC-end

Department of Vegetable Crops, R.W. Michelmore
University of California at Davis (UCD)
Assumundson Hall, UCD, Davis, CA 95616, USA
Tel: 1 (530) -742-1742
Fax: 1 (530) -752-9659
Email: akozik@atcg.org [michelmorevegmail.ucdavis.edu]
Belongs to contig_QG_Cantig656, see <http://cgpdb.ucdavis.edu/>
Plate: QCD9s row: A col: 22

http://www.edg.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billand at CEPR (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

Location/Qualifiers
1. 1101
/orientem-ubccat11

ganjiam="Droscob:1

FEATURES	LOW: A	COLUMN: 23
source	Location/Qualifiers	
	1	745

```

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/ab_initio="1"

```

type="genomic DNA"

```

/cname="BACN37f1.0"
/clone_lib="DrosBAC"
/plasmid="pbeloBAC11"
/note="end : Sp6"

```

```
ce="end : SP6"
```

```
/lab_host="E.coli"  
/clone_id="QG_ABCDI lettuce salinas"  
/note="vector: pBRCDMS1AB; The library was constructed  
from 10 different sources of RNA from a single genotype.
```

Query Match	3.6%;	Score 46.2;	DB 9;	Length 1101;
Best Local Similarity	21.8%;	Pred. No. 0.95;	Matches 83;	Conservative 126;
			Mismatches 172;	Indels 0;
			Gaps	

2222222222222222

Db	1007	KTRITKMKMWRARBAKMRDWMWDTPMDMDTKMKDDRTKMTATRGDBRAGRDRDR	948
Oy	615	GATTGTATATAGAGAATTTACGCCAGCTTGTATTAATGACGTGTTTCAGTTCTAGAAG	674
Db	947	RGDRRRRRRRRTTRKRWRRRAAMWTAMTWMWGTTKGAMWTKTRKRKGGGAAAADWGR	888
Oy	675	TGATCTCGAGAAGTCCCCACAGACTGTGAAAAATTTGATTGAAGATCACTGGAAATTC	734
Db	887	TDDWMDRRAAAAAGKDGKGGKATGWAAGAAAAAAMDYKTAATKTKRKTGTGRAMWAAG	828
Oy	735	ACATGAAGATGTACATTACATACGSAATCAAGTTATTAATGGGGACGTCAGAGGTG	794
Db	827	TRTTTDAATATMTKRAAAAAGRRAAAAMKDRGTGRKGRKKRGDTPDGCTKMTRTTWT	768
Oy	795	CTCAGAAAACCCCAATTGTTATCAAGACTACTCCAAATGAAAAAGAAAGCCAGAGA	854
Db	767	KTAARARARARAAAAGMTKXTDGRARAAWMTTTRTKGAATDRGAAAAAATTGGA	708
Oy	855	AAGCTCTGTTGTGTATGACAGCTTAATAGTCCAAATTTACAGTTATCAAGAAATG	914
Db	707	AAAAAGGTTTDAAGAAATAAAMWRRTTWDWMAAARAAGAAWAAKAAADAADAKKGG	648
Oy	915	GGAATGCGTGAAAAATACGA	935
Db	647	RRAKRAAATAARRAKKAKGA	627
RESULT 31			
CG067021			
LOCUS	PUIDB42TD ZM.0.6.1.0 KB	Zea mays genomic clone ZM87A055H12,	
DEFINITION	genomic survey sequence.		
ACCESSION	CG067021		
VERSION	CG067021.1	GI:3339201	
KEYWORDS	GSS.		
SOURCE			
ORGANISM	Zea mays		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 727)		
AUTHORS	WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Benneetzen,J.		
TITLE	Maize Genomic Consortium		
JOURNAL	Unpublished (2003)		
COMMENT	Other_GSSs: PUIDB42TB Contact: Cathy WhiteLaw		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA		
	Tel: 301-838-5843		
	Fax: 301-838-0208		
	Email: whiteLaw@cigr.org		
	Seq primer: TP		
	Class: sheared ends.		
FEATURES			
source	1..727	Location/Qualifiers	
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/strain="B73"		
	/db_xref="taxon:4577"		
	/clone="ZM87A055H12"		
	/clone.lib="ZM.0.6.1.0 KB"		
	/note="Vector: PCR4-TOPO, Site 1: EcoRI; 0.6-1.0 kb high cot selected genomic DNA library"		
ORIGIN			
Query Match	3.5%;	Score 45.8;	DB 9; Length 727;
Best Local Similarity	53.7%;	Pred. No. 1.1;	
Matches	95;	Conservative 0;	Mismatches 82; Indels 0; Gaps 0
464	CAGTTGTTGCCAGAGTTATTATCCATTATATCCCAAGTATTCAGTAGCAATTAAG	523	

Db	Query Match	Score	DB	Length	bp
444	CTGTAACTCTCATGTTTCAGGAACATCTTCTGGCGACATTCGAAAGTGAAGCAACTTG	37.5%	Pred. No. 1.2;	499	503
Qy	524 ATATTTCGACACAGTTTCACAGATTGCACATGTTATGCCCCAAAAAGATAGATCTG	3.5%	Score 45.4;	DB 9;	583
Db	504 ACTTGTTTCACGCTATTTGCGCCAAAGTGCACATGTTTATGACACAAAGGAAATTTGGCAGT				563
Qy	584 GCTTGATNGTTCGACATCGCAATTTATATGCTGATTTGATATGATGAAAGATTTTCAGCAG				640
Db	564 GTTTTGATGTTAGTGTCTGTCTTATGAGAGTCACCGCTATGTTAAGGTTTCTTCAG				620
RESULT 32					
CNS06D75					
LOCUS					
DEFINITION	CNS06D75	499 bp	DNA	linear	GenScope
ACCESSION	T7 end of clone AROAA011A09 of library AROAA from strain CBS 732 of				
VERSION	Zygosaccharomyces rouxii, genomic survey sequence.				
KEYWORDS	AL393447				
ORGANISM	AL393447.1	GI:12143650			
SOURCE	GenScope.				
REFERENCE	Zygosaccharomyces rouxii				
AUTHORS	Zygosaccharomyces rouxii				
	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
	Yeast; Fungi; Ascomycota; Saccharomycetaceae; Zygosaccharomyces.				
	1 (bases 1 to 499)				
	Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,				
	Boletini-Pukhara,M., Bon,E., Broctier,P., Caaserego,S.,				
	de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Florente,B.,				
	Malpêtre,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,				
	Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,				
	Wincker,P. and Weissenbach,J.				
	Genomic exploration of the Hemiascomycetous yeasts: 1. A set of				
	yeast species for molecular evolution studies				
	FEBS Lett. 487 (1), 3-12 (2000)				
JOURNAL	20584711				
MEDLINE	1152876				
PUBMED	2 (bases 1 to 499)				
REFERENCE	de Montigny,J., Straub,M., Potier,S., Tekala,F., Dujon,B.,				
AUTHORS	Wincker,P., Artiguenave,F. and Souciet,J.				
	Genomic exploration of the hemiascomycetous yeasts: 8.				
	Zygosaccharomyces rouxii				
	FEBS Lett. 487 (1), 52-55 (2000)				
JOURNAL	20584718				
MEDLINE	1152883				
PUBMED	3 (bases 1 to 499)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
	Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage,				
	2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :				
	beiret@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
	This GSS is part of a random genomic sequencing program of thirteen				
	yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces				
	exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,				
	Saccharomyces kluyveri, Kluyveromyces thermocolerans, Kluyveromyces				
	lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia				
	angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,				
	Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to				
	5 kb were prepared and both extremities were sequenced. See				
	keywords for description of this sequence and for the sequence of				
	the other extremity of this insert.				
	Location/Qualifiers				
	1. 499				
	/organism="Zygosaccharomyces rouxii"				
	/mol_type="genomic DNA"				
	/strain="CBS 732"				
	/db_xref="taxon:4956"				
	/clone="AROAA011A09"				
	/clone_id="AROAA"				
	/note="end : T7"				
FEATURES					
SOURCE					
ORIGIN					
Query Match	3.5%	Score 45.4;	DB 9;	Length 499;	
Best Local Similarity	37.5%;	Pred. No. 1.2;			

Query Match	Best Local Similarity	Score	DB	Length
Matches 771	Conservative 0	Mismatches 54	Indels 0	Gaps 0
517	AATTAAGATATTTTGACACAGTTGCACAGATTGCACATTGTTATGCCCAAAAAAGATA	576		
562	AACTTGAATGATGTCATATTAATACCCCAANTGGCACACGTAATTCCTCAAGAAAGATT	621		
577	GGATCGGGTTTGATGATTCGCACTGCACATTAATGCTGATTTGATATATAGAAATTTGAG	636		
622	GGGAGTGATTTGAATGTTAGTTGACGTGTTATGGAAGCATCGTTATGTAAAGATTTC	681		
637	CCAGCTTTGAT	647		
682	CCACAAGTTAT	692		

RESULT	34
BMI65227/c	
LOCUS	
DEFINITION	BMI65227 697 bp mRNA
LocusTag	EST56750 PyBS Plasmodium yoelii yoe1l cDNA clone PYCMJ47 5' end,
ACCESSION	mRNA sequence.
VERSION	BMI65227
KEYWORDS	BMI65227.1 GI:17310908
SOURCE	EST.
	Plasmodium yoelii yoe1l

REFERENCE	1 (baes 1 to 697)
AUTHORS	Carlton,J.M., Daly,T.M., Long,C.A., Bergman,L.W., Valdya,A.B.
TITLE	Frazer,C.M. and Carucci,D.J. Plasmodium yoelii: EST products of the Plasmodium yoelii: EST products of the

TITLE Plasmodium yoelii EST project at TIGR
JOURNAL Unpublished (2001)
COMMENT Contact: Jane Carlton
Parasite Genomics Group
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-550-8919
Fax: 301-858-2519

```

FEATURES
    source      Location/Qualifiers
    1..697
    /organism="Plasmodium yoelii yoelii"
    /mol_type="mRNA"
    /strain="17XI"
    /sub_species="yoelii"

```

collected from BALB/cByJ mice 20-25% parasitemia, blood was parasitized, and leukocytes removed by passage over microcrystalline cellulose columns. Total RNA was isolated using the guanidinium isothiocyanate method, and mRNA isolated using oligo(dT)-cellulose chromatography. First strand cDNA synthesis was completed using a 50-base primer and reverse transcriptase in the presence of 5-methyl dCTP. After second strand synthesis, eleven cDNA termini were treated with Pfu DNA polymerase and EcoRI adaptors ligated to the blunt ends. The sample was ligated with XhoI and separated on a Sephacryl S-500 column. Size-fractionated cDNA was precipitated and ligated to

melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR2703"
/clone_11b="RPc1-98"
/note="end : TET3"

ORIGIN

Query Match 3.4%; Score 44.8; DB 9; Length 1101;
Best Local Similarity 28.6%; Pred. No. 2.2; Mismatches 0; Gaps 0;
Matches 166; Conservative 111; Indels 0; Gaps 0;

699 GTTGAATAATTTGATGAAAGTAACTGGAATTCGAACGATGATGATACCATTA 758
1081 GKAAATWTAATRGMAARAATTTTGTGAAAAATRTWAMDKMWRRAAAAKTDK 1022
759 CGAATCAAGTTATTAATGAGTCAAGGCTGCTCAGAAACACCAATTTGATTC 818
1021 TTDAATAAAWMAATTAATAATGTTTMTAAARATWAAAAAATAAKTRMAWR 962
819 ACGAGTACTCCAATGAGAAAAAGCAAGAAAGCTGTTGTTGATGACCACT 878
961 RGGGAGATTGAGWAAAAAGARAAAAAARRRGAAGTGAAGGTGGAARGAATTG 902
879 TAATAGTCCAAATTTACAGTTTATGAGGAATTTGAGGAATGCGTGAATAACGATC 938
901 KTAAGRGATTAATAATTTTSMAAAAAVDTSTSAASAAAAVAAAAA 842
939 AGACCCAGAGCTTATATTAAGAGTTAGTATCTGTTGACCTTTGACCTTGCAT 998
841 AAAAAAAWMAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 782
999 TAAAGACATCAAGAAAGGTTCAAGCATTAACAAATAACAGCTTCCATTAAC 1058
781 RARSARGAARATWAAAAAAGAAAAAASVSAASVAMRAVSSASAAAGAAWAA 722
1059 TGATGTCCAAACCCAGTTGTGACCGTTGTCAGAGATTCTGTTGTTGTTGTTGT 1118
721 ARVAAAATTTGATTAAGGTGGAGAAAAATTTWAAATTAAGMWRRAATGTAARARAT 662
1119 GGTTCACAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1178
661 ATGAARSAGATTAATTAATAATAATAATAATAATAATAATAATAATAATAATA 602
1179 AATATTTAAGCAAGAACTCTTTGAAATCCAGATTTTCTAATGTTTACTGCTTGA 1238
601 AAVTSAASASASASTAASWMAASATSSAAASATTAATAATAATAATAATAATA 542
1239 TTGGAAGACCAAGAGGTGATCTGTAAGAAACCA 1278
541 AVSSASAAATWTAATAATAATAATAATAATAATAATAATAATAATAATAATA 502

RESULT 39
AZ549980 900 bp DNA linear GSS 14-NOV-2000
LOCUS ENTDD94TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, genomic survey sequence.
ACCESSION AZ549980

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ549980.1 GI:11175122
GSS.
Entamoeba histolytica
Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
1 (bases 1 to 900)
Loftus, B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjoftus@igf.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 20
High quality sequence stop: 890.
Location/Qualifiers
1..900
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/clone_11b="Entamoeba histolytica Sheared DNA"
/note="Vector: pHO51, Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The V + I method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaubin and B.
Barell, Oxford University Press, 1999)."

FEATURES

source

1027 TTAACACAAAATATGAGAGTTCACATTAACCTGATGCCAACCCAGTTGTGACCGT 1086
307 TTGAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 366
1087 TGTCAAGAGATTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1146
367 GATGAAGCGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGAT 426
1147 ATAGCTATTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1206
427 GATGATGAAGAGATGAAGAGATGAAGAGATGAAGAGATGAAGAGATGAAGAT 486
1207 CCAGATTTTTCATTAATGTTTACTGCGTTGATTTGGAAGCAACCAAGAGTACTT 1266
487 GAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGAT 546
547 GATGAAGATGATGAAGATGATGAAGAT 573

ORIGIN

Query Match 3.4%; Score 44.6; DB 8; Length 900;
Best Local Similarity 47.9%; Pred. No. 2.4;
Matches 128; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

RESULT 40
CNS00E07 1101 bp DNA linear GSS 04-JUN-1999
LOCUS CNS00E07

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 03:01:13 ; Search time 483 Seconds
(without alignments)
14118.008 Million cell updates/sec

Title: US-10-069-062-6
Perfect score: 1299
Sequence: 1 atgcacaaagcatttcgtc.....aagactataggtttataa 1299

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23sep04:*
1: genebegn19808:*
2: genebegn19908:*
3: genebegn20008:*
4: genebegn20018:*
5: genebegn20028:*
6: genebegn20038:*
7: genebegn20048:*
8: genebegn20058:*
9: genebegn20068:*
10: genebegn20078:*
11: genebegn20088:*
12: genebegn20098:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1299	100.0	1299	4 AAD02792	Aad02792 Candida a
2	1299	100.0	1763	4 AAD02791	Aad02791 Candida a
3	1286.2	99.0	1299	4 AAS23422	Aas23422 Candida a
4	1286.2	99.0	1299	6 ABZ31736	Abz31736 Candida a
5	521	40.1	577	4 AAD02788	Aad02788 S. cerevi
6	470.4	36.2	547	4 AAD02787	Aad02787 S. cerevi
7	132	10.2	1356	6 AAL40805	Aal40805 Nucleic a
8	132	10.2	1356	6 AAD31013	Aad31013 Yeast pho
9	132	10.2	1356	6 ABK96801	Abk96801 S. cerevi
10	132	10.2	1356	12 ADF23287	Adf23287 MevB oper
11	132	10.2	1356	12 ADI82315	Adi82315 S. cerevis
12	132	10.2	1356	12 ADI14834	Adi14834 Yeast pho
13	132	10.2	4482	12 ADF23291	Adf23291 Mevalonat
14	132	10.2	4482	12 ADI82319	Adi82319 Mevalonat
15	132	10.2	5051	12 ADF23294	Adf23294 Recombina
16	132	10.2	5051	12 ADI82322	Adi82322 Amorphin-4
17	132	10.2	5963	12 ADF23295	Adf23295 Recombina
18	132	10.2	5963	12 ADI82323	Adi82323 Amorphin-4
19	132	10.2	7681	6 AAD31026	Aad31026 Operon D
20	132	10.2	7681	6 AAD31023	Aad31023 Operon A
21	132	10.2	7695	6 AAD31024	Aad31024 Operon B

22	132	10.2	8224	6 AAD31027	Aad31027 Operon E
23	132	10.2	8235	6 AAD31025	Aad31025 Operon C
24	132	10.2	8400	6 AAD31029	Aad31029 Operon G
25	132	10.2	9253	12 ADF23289	Adf23289 Isopenen
26	132	10.2	9253	12 ADI82317	Adi82317 Amorphin-4
27	132	10.2	13917	6 AAD31037	Aad31037 Plasmid c
28	132	10.2	14623	6 AAD31039	Aad31039 Plasmid c
29	132	10.2	14623	6 AAD31041	Aad31041 Plasmid c
30	67.6	5.2	1506	10 ADB69813	Adb69813 C. neofo
31	66	5.1	1356	8 ABT19226	Abt19226 Aspergill
32	65	5.0	65	6 ABZ28940	Abz28940 Candida g
33	65	5.0	90	4 AAS23667	Aas23667 Tetraacycl
34	62.6	4.8	1455	8 ABT21056	Abt21056 Aspergill
35	62.6	4.8	1509	8 ABT20458	Abt20458 Aspergill
36	62.6	4.8	1509	8 ABT18642	Abt18642 Aspergill
37	62.6	4.8	3508	8 ABT18048	Abt18048 Aspergill
38	62.6	4.8	3509	8 ABT19862	Abt19862 Aspergill
39	53.2	4.1	869	8 ABZ51452	Abz51452 Aspergill
40	48.2	3.7	2000	8 ADA71938	Ada71938 Rice gene
41	46.8	3.6	586	3 AAF08722	Aaf08722 Fusarium
42	45	3.5	2000	8 ADA71938	Ada71938 Rice gene
43	44.6	3.4	1356	12 ADL03441	Adl03441 DNA encod
44	43	3.3	1341	8 ACA39227	Ac39227 Prokaryot
45	43	3.3	119211	4 AAF28553	Aaf28553 Genomic F

ALIGNMENTS

RESULT 1	
ID AAD02792	standard; DNA; 1299 BP.
XX	
AC AAD02792;	
XX	
DT 31-MAY-2001	(first entry)
XX	
DE Candida albicans phosphomevalonate kinase (PMK) coding sequence.	
XX	
KW Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;	
XX	
OS Candida albicans.	
XX	
FM Key	Location/Qualifiers
FT CDS	1..1299
FT	/*tag= a
FT	/product= "C. albicans phosphomevalonate kinase (PMK)"
FT	/transl_except= (pos:730..732, aa:Glu)
XX	
PN MO200114533-A2.	
XX	
PD 01-MAR-2001.	
XX	
PF 15-AUG-2000; 2000MO-GB003100.	
XX	
PR 21-AUG-1999; 99GB-00019766.	
XX	
PA (ASTR) ASTRAZENECA AB.	
XX	
PA (ASTR) ASTRAZENECA UK LTD.	
XX	
PI Rosamond JDC, Schnell NF;	
XX	
DR WPI, 2001-218441/22.	
XX	
DR P-PSDB; AAY72679.	
XX	
PT New polypeptides and polynucleotides (ERG8) from Candida albicans, useful	
XX	
PT in assays for identifying inhibitors of phosphomevalonate kinase activity	
XX	
PS and as reagents for diagnosing C. albicans infection.	
XX	
PS Claim 6; Page 26; 29pp; English.	
XX	
XX The present sequence is phosphomevalonate kinase (PMK; ERG8) coding	

CC sequence from *Candida albicans*. The ERG8 protein is useful in an assay
 CC for identifying compounds that inhibit phosphomevalonate kinase (PMK)
 CC activity. These inhibitors are useful as anti-fungal agents. The ERG8 DNA
 CC and protein are also useful as reagents for diagnosing *C. albicans*
 CC infection

XX
 SQ Sequence 1299 BP; 439 A; 220 C; 273 G; 367 T; 0 U; 0 Other;

Query Match 100.0%; Score 1299; DB 4; Length 1299;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGTCAAAAGCATTAGTGCACCTGGAAAAGCATTTCTTGCTGGTGAATATTTGGTCTT 60
DB 1 ATGTCAAAAGCATTAGTGCACCTGGAAAAGCATTTCTTGCTGGTGAATATTTGGTCTT 60
QY 61 GAGCCAAATTTAGATGCTATGAGACAGATTTGTCATCAGAAATGATGCACTTATTAACA 120
DB 61 GAGCCAAATTTAGATGCTATGAGACAGATTTGTCATCAGAAATGATGCACTTATTAACA 120
QY 121 CCAAAAGAACCGATTGAAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGCAAC 180
DB 121 CCAAAAGAACCGATTGAAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGCAAC 180
QY 181 GGAAGATGGGAATATCATATCATCAATCAATACAGAGAACCCAGAGAAATTCAGTCACGC 240
DB 181 GGAAGATGGGAATATCATATCATCAATCAATACAGAGAACCCAGAGAAATTCAGTCACGC 240
QY 241 ATAAATCCATTTTATAGGCACTATATCATCGTTTATGCTTATATTTCAACCCGCA 300
DB 241 ATAAATCCATTTTATAGGCACTATATCATCGTTTATGCTTATATTTCAACCCGCA 300
QY 301 GCATTTGATCTTGAATCATCATTTTACTCAGACCCCTGATATCATTCACAGAAATATCT 360
DB 301 GCATTTGATCTTGAATCATCATTTTACTCAGACCCCTGATATCATTCACAGAAATATCT 360
QY 361 GAAACCAAGACATCTCCGAAATGAGAAAAACATTTCTTTACCATTTCTGTCCTAAC 420
DB 361 GAAACCAAGACATCTCCGAAATGAGAAAAACATTTCTTTACCATTTCTGTCCTAAC 420
QY 421 GAAGTGAAGAAACCGGATTAGGTTTCATGCGAGAGATTAGTGCAGTTGGCCCAAGT 480
DB 421 GAAGTGAAGAAACCGGATTAGGTTTCATGCGAGAGATTAGTGCAGTTGGCCCAAGT 480
QY 481 TTAATATCCATTTTATCCCAATGTTTACAGTAAGAAATATTTTGGCACAACCT 540
DB 481 TTAATATCCATTTTATCCCAATGTTTACAGTAAGAAATATTTTGGCACAACCT 540
QY 541 GACAGATTTGACATTTGTTAGCCCAAAAAAGATAGATGATGGGTTGATGTTGCAACT 600
DB 541 GACAGATTTGACATTTGTTAGCCCAAAAAAGATAGATGATGGGTTGATGTTGCAACT 600
QY 601 GCAATTTATGCTCTGATTTGATATAGAAATTTGATTCAGCAAGCTTTGATTAATGAGT 660
DB 601 GCAATTTATGCTCTGATTTGATATAGAAATTTGATTCAGCAAGCTTTGATTAATGAGT 660
QY 661 CAGGTTCTAGAAAGTGAATCTCGAAGAGTTCCCAAGAGTTGAAAAATTTGATGAAAGT 720
DB 661 CAGGTTCTAGAAAGTGAATCTCGAAGAGTTCCCAAGAGTTGAAAAATTTGATGAAAGT 720
QY 721 AACTGGGAATTCAAACATGAAAGATGATCAATTCATACGGAATTAAGTTATATGAGT 780
DB 721 AACTGGGAATTCAAACATGAAAGATGATCAATTCATACGGAATTAAGTTATATGAGT 780
QY 781 GACGTCAGAGGTGCTCAGAAAACACCAAAATGATACAGAGTCTCAATGGAAG 840
DB 781 GACGTCAGAGGTGCTCAGAAAACACCAAAATGATACAGAGTCTCAATGGAAG 840
QY 841 GAAAAGCCAGAAAGAGCTCTGTTGTATGACAGCTTAATATGAGCCAAATTTACAGTTT 900
DB 841 GAAAAGCCAGAAAGAGCTCTGTTGTATGACAGCTTAATATGAGCCAAATTTACAGTTT 900
QY 901 ATGAAGGAATGAGGGAATGCTGAAAAAATACGACTCAGACCCAGAGACTTATATTTAA 960

```

```

DB 901 ATGAAGGAATGAGGGAATGCTGAAAAAATACGACTCAGACCCAGAGACTTATATTTAA 960
QY 961 GAGTTAGATCATTTCTGTTGAGCCTTTGACCTGTGCGATTAAGAACATCAGAAAAGGTTA 1020
DB 961 GAGTTAGATCATTTCTGTTGAGCCTTTGACCTGTGCGATTAAGAACATCAGAAAAGGTTA 1020
QY 1021 CAAGCATTAACCAAAAATCAGAGTTTCCATTTGAACTGTATGTCCAAACCCAGTTGTTG 1080
DB 1021 CAAGCATTAACCAAAAATCAGAGTTTCCATTTGAACTGTATGTCCAAACCCAGTTGTTG 1080
QY 1081 GACCGTTGTCAGAGATTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
DB 1081 GACCGTTGTCAGAGATTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
QY 1141 GATGCAATAGCTGATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
DB 1141 GATGCAATAGCTGATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
QY 1201 GAAATCCAGATTTATTTTATATGTTTACTGCGGTGATTTGGAAGCAACAGAGGT 1260
DB 1201 GAAATCCAGATTTATTTTATATGTTTACTGCGGTGATTTGGAAGCAACAGAGGT 1260
QY 1261 GTACTTGAAGAAAAACAGAGACTATATAGTTTATTA 1299
DB 1261 GTACTTGAAGAAAAACAGAGACTATATAGTTTATTA 1299

```

RESULT 2
 AAD02791
 ID AAD02791 standard; DNA; 1763 BP.

XX AAD02791;
 AC
 XX
 DT 31-MAY-2001 (first entry)

DE *Candida albicans* phosphomevalonate kinase (PMK) gene (ERG8 gene).
 XX
 KW Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;
 XX infection; de.
 OS *Candida albicans*.

XX
 FH Key Location/Qualifiers
 FT CDS 89..1387
 FT /tag= a
 FT /product= "C. albicans phosphomevalonate kinase (PMK)"
 FT /trans_except= (pos:818..820, aa:Glu)

XX MO200114533-A2.

XX 01-MAR-2001.

XX 15-AUG-2000; 2000WC-GB003100.

XX 21-AUG-1999; 99GB-00019766.

XX (ASTR) ASTRAZENECA AB.
 PA (ASTR) ASTRAZENECA UK LTD.

XX Rosamond JDC, Schnell NF;

XX WPI; 2001-218441/22.

XX P-PSDB; AAY72679.

XX New polypeptides and polynucleotides (ERG8) from *Candida albicans*, useful
 PT in assays for identifying inhibitors of phosphomevalonate kinase activity
 PT and as reagents for diagnosing *C. albicans* infection.
 XX
 PS
 XX

XX Claim 6; Fig 1; 29pp; English.

CC The present sequence is phosphomevalonate kinase (PMK) gene (ERG8 gene)
 CC from *Candida albicans*. The ERG8 protein is useful in an assay for

CC identifying compounds that inhibit phosphomethyltransferase kinase (PMK)
CC activity. These inhibitors are useful as anti-fungal agents. The ERG8 DNA
CC and protein are also useful as reagents for diagnosing *C. albicans*
CC infection

SQ Sequence 1763 BP; 582 A; 300 C; 349 G; 530 T; 0 U; 2 Other;

Query Match	100.0%	Score 1299	DB 4	Length 1763
Best Local Similarity	100.0%	Pred. No.	0	
Matches 1299	0	Mismatches	0	Gaps 0

Qy	1	ATGTCAAAAGCATTTGTGTCACCTGCGAAAAGCATTTCTTGCGGGGATATTTGGTCTT	60
Db	89	ATGTCAAAAGCATTTGTGTCACCTGCGAAAAGCATTTCTTGCGGGGATATTTGGTCTT	148
Qy	61	GAGCCAAATTTATGATGCTTATGTGACAGCATTTGTCAACGAATGATGACGTTATPACA	120
Db	149	GAGCCAAATTTATGATGCTTATGTGACAGCATTTGTCAACGAATGATGACGTTATPACA	208
Qy	121	CGAAAAGGAACGAGTTTGGAAAGAAATCTGAATTCAAAATTTCTTCACCCCAATTTGGAAC	180
Db	209	CGAAAAGGAACGAGTTTGGAAAGAAATCTGAATTCAAAATTTCTTCACCCCAATTTGGAAC	268
Qy	181	GGAGAATGGGAATATCACTATTCATCAATAACAGAGAAGCCACAGAGAGTTCACTACAGC	240
Db	269	GGAGAATGGGAATATCACTATTCATCAATAACAGAGAAGCCACAGAGAGTTCACTACAGC	328
Qy	241	ATPAAATCCATTTTAAAGGCAACTATATTCAATGTTTTAGCTTATATTCAACCGACGAA	300
Db	329	ATPAAATCCATTTTAAAGGCAACTATATTCAATGTTTTAGCTTATATTCAACCGACGAA	388
Qy	301	GCATTTGATCTTGAATTCATCAATTTACTACAGACCTGGAATPACTTCAACAGAAATACT	360
Db	389	GCATTTGATCTTGAATTCATCAATTTACTACAGACCTGGAATPACTTCAACAGAAATACT	448
Qy	361	GAACCAAGACATCTCGAATGGAAGAAAAACATTTCTTAAACATTCGTGTCATTACC	420
Db	449	GAACCAAGACATCTCGAATGGAAGAAAAACATTTCTTAAACATTCGTGTCATTACC	508
Qy	421	GAAGTGGAAAAAGCCGATTTAGGTTTCATGCGAGATTAGTGCAGTGTGTCACAAAGT	480
Db	509	GAAGTGGAAAAAGCCGATTTAGGTTTCATGCGAGATTAGTGCAGTGTGTCACAAAGT	568
Qy	481	TTATTTATCCCATTTTATCCCAATGTTTACATAGGAATPAAAGATTTTGCACAAGTT	540
Db	569	TTATTTATCCCATTTTATCCCAATGTTTACATAGGAATPAAAGATTTTGCACAAGTT	628
Qy	541	GCAAGATTGACATTTGTTATGCGCCAAAAAAAGATAGATCTGGGTTGATGTTGCAACT	600
Db	629	GCAAGATTGACATTTGTTATGCGCCAAAAAAAGATAGATCTGGGTTGATGTTGCAACT	688
Qy	601	GCAATTTTATGCTGATTTGTATATAGAAATTTTCACGACGTTTGATTAAGAAGCTGTT	660
Db	689	GCAATTTTATGCTGATTTGTATATAGAAATTTTCACGACGTTTGATTAAGAAGCTGTT	748
Qy	661	CAGGTTCTAGAAGATATCCTGAGAGTTCCCAAGAGTTGAAAAAATGATTGAAAGT	720
Db	749	CAGGTTCTAGAAGATATCCTGAGAGTTCCCAAGAGTTGAAAAAATGATTGAAAGT	808
Qy	721	AACTGGGAATTCAAACATGAAGATGTACATTACATACGGAATCAAGTTATTAAGGCT	780
Db	809	AACTGGGAATTCAAACATGAAGATGTACATTACATACGGAATCAAGTTATTAAGGCT	868
Qy	781	GAGCTCAAGGGTGGCTCAGAAACACCCAAATTTGATACAGATCTCCATATGAAAAAG	840
Db	869	GAGCTCAAGGGTGGCTCAGAAACACCCAAATTTGATACAGATCTCCATATGAAAAAG	928
Qy	841	GAAGAAGCAGAGAAAGCTCTGTGTGTATGACAGCTTAATATGTGCCAATTTTACAGTTT	900
Db	929	GAAGAAGCAGAGAAAGCTCTGTGTGTATGACAGCTTAATATGTGCCAATTTTACAGTTT	988
Qy	901	ATGAAAGAAATTGAGGGAATTCGTGAAAAATTCGACTCAGACCCGAGACCTTATATTAAA	960

[illegible]

RESULT 3
 AAS23422
 ID AAS23422 standard; DNA; 1299 BP.
 AC AAS23422;
 DT 04-DEC-2001 (first entry)
 DE Candida albicans essential gene CAYMR220W (ERG8).
 KW Gene identification; essential gene; GRACE; pathogenic fungus;
 KM gene replacement and conditional expression; fungal infection; ds.
 OS Candida albicans.
 PN WO200160975-A2.
 PP 23-AUG-2001.
 PE 20-FEB-2001; 2001WO-US005551.
 PR 18-FEB-2000; 2000US-0183534P.
 PX (ELIT-) ELITRA PHARM INC.
 PI Roemer T, Jiang B, Boone C, Bussey H;
 DR WPI; 2001-489080/53.
 DR P-PSDB; AAU15094.
 PT Identifying genes essential to fungal metabolisms and identifying
 PT potential therapeutic agents that target these genes.
 PS Claim 22; Page 162; 324pp; English.
 CC The present invention relates to novel methods for constructing fungal
 CC strains useful for identification and validation of gene products as
 CC targets for therapeutic agents, for creating a collection of identified
 CC essential genes, and screening assays for the discovery of new drugs. The
 CC invention provides the GRACE (gene replacement and conditional
 CC expression) method for the construction of mutant organisms referred to
 CC as GRACE strains of the organism. The invention can be applied to any
 CC organism, particularly a pathogenic fungus e.g. Candida albicans,
 CC Aspergillus fumigatus and Cryptococcus neoformans. The methods are useful
 CC to identify agents that may be used in the treatment of fungal
 CC infections. AAS23381-AAS23442 represent C. albicans essential genes

CC that contributes to the resistance of a diploid fungus to an antifungal
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus
 CC and for identifying a therapeutic agent for treatment of a mammalian
 CC disease. (M1) is useful for identifying a compound which modulates the
 CC activity of a gene product, preferably enzymatic activity, carbon
 CC compound catabolism, biosynthetic, transporter, transcriptional,
 CC translational, signal transduction, DNA replication and cell division
 CC activity. The method is useful for identifying a compound having the
 CC ability to inhibit growth or proliferation of *C. albicans* cells and for
 CC treating infection by *C. albicans*. The present sequence is that of an
 CC essential *Candida albicans* gene used in the method of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office
 CC XX

Sequence 1299 BP; 439 A; 221 C; 271 G; 368 T; 0 U; 0 Other;

Query Match 99.0%; Score 1286.2; DB 6; Length 1299;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AAGTAAAGCAATTTAGTACCTGGAAAGCAATTTCTGCTGGATATTTGTTCTT 60
 DB 1 AAGTAAAGCAATTTAGTACCTGGAAAGCAATTTCTGCTGGATATTTGTTCTT 60
 QY 61 GAGCAATTTATGATGCTTATGATGACGATTTGTCATCAAGATGATGATTTATACA 120
 DB 61 GAGCAATTTATGATGCTTATGATGACGATTTGTCATCAAGATGATGATTTATACA 120
 QY 121 CCAAAAGAACCAATTTGAAAGATCTAGATCAAAATTTCTTACCCCAATTTGCAAC 180
 DB 121 CCAAAAGAACCAATTTGAAAGATCTAGATCAAAATTTCTTACCCCAATTTGCAAC 180
 QY 121 CGAAGATGGGATATTCATATCATATCAAGATCAAGATCAAGATCAAGATCAAG 240
 DB 121 CGAAGATGGGATATTCATATCATATCAAGATCAAGATCAAGATCAAGATCAAG 240
 QY 181 GGAAGATGGGATATTCATATCATATCAAGATCAAGATCAAGATCAAGATCAAG 240
 DB 181 GGAAGATGGGATATTCATATCATATCAAGATCAAGATCAAGATCAAGATCAAG 240
 QY 241 ATAAATCCATTTTATGAGGCAATATATTCATGCTTTATGCTTATATTCACACCGAA 300
 DB 241 ATAAATCCATTTTATGAGGCAATATATTCATGCTTTATGCTTATATTCACACCGAA 300
 QY 301 GCATTTGATCTTGAATCATATTTATCTGCAAGATTTGCTTATATTCACACCGAA 360
 DB 301 GCATTTGATCTTGAATCATATTTATCTGCAAGATTTGCTTATATTCACACCGAA 360
 QY 361 GAAACCAAGACATCTCGAATGAGAAACATTTCTTATCAATCTGCTGATTAAC 420
 DB 361 GAAACCAAGACATCTCGAATGAGAAACATTTCTTATCAATCTGCTGATTAAC 420
 QY 421 GAAGTGAAGAACCGGATTTAGGTTATGCTGAGATTTGCTGATTTGCTGCAAGT 480
 DB 421 GAAGTGAAGAACCGGATTTAGGTTATGCTGAGATTTGCTGATTTGCTGCAAGT 480
 QY 481 TTATATCCCATTTATCCCAATGTTATCAAGATTAAGATATTTGCAACGTT 540
 DB 481 TTATATCCCATTTATCCCAATGTTATCAAGATTAAGATATTTGCAACGTT 540
 QY 541 GACAGATTTGACATTTGTTATGCTGCAAGATTTGCTGATTTGCTGCAAGT 600
 DB 541 GACAGATTTGACATTTGTTATGCTGCAAGATTTGCTGATTTGCTGCAAGT 600
 QY 601 GCAATTTATGCTGATTTGTTATGAGATTTGCTGCAAGATTTGCTGCAAGT 660
 DB 601 GCAATTTATGCTGATTTGTTATGAGATTTGCTGCAAGATTTGCTGCAAGT 660
 QY 661 CAGGTTCTGAGAAAGTATCTGGAAGTCCCAAGAGTGAAGATTTGATTTGCAAGT 720
 DB 661 CAGGTTCTGAGAAAGTATCTGGAAGTCCCAAGAGTGAAGATTTGATTTGCAAGT 720
 QY 721 AACTGGGAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 DB 721 AACTGGGAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 781 GACGTCAAGGCTGCTCAGAAACCCCAATTTGATTCAGGATCTCCATATGAAAAAG 840
 DB 781 GACGTCAAGGCTGCTCAGAAACCCCAATTTGATTCAGGATCTCCATATGAAAAAG 840
 QY 841 GAAAGCCAGAAAGAGCTCTGTTGATGACAGCTTAAATGAGCCAAATTTACGTT 900
 DB 841 GAAAGCCAGAAAGAGCTCTGTTGATGACAGCTTAAATGAGCCAAATTTACGTT 900
 QY 901 ATGAAGATTTAGGGAATGCGTGAATAATGCACTCAGACCCAGACCTTATATTA 960
 DB 901 ATGAAGATTTAGGGAATGCGTGAATAATGCACTCAGACCCAGACCTTATATTA 960
 QY 961 GAGTTAGATCATTTCTGTTGAGCTTTGACTGTTGCGATTAAGACATCAAGAAAGGTTA 1020
 DB 961 GAGTTAGATCATTTCTGTTGAGCTTTGACTGTTGCGATTAAGACATCAAGAAAGGTTA 1020
 QY 1021 CAGCATTTAACAAGAAATCAAGATTCATATTTGAACCTGATGCAACCAAGTTGTTG 1080
 DB 1021 CAGCATTTAACAAGAAATCAAGATTCATATTTGAACCTGATGCAACCAAGTTGTTG 1080
 QY 1081 GACCGTTGCAAGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
 DB 1081 GACCGTTGCAAGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
 QY 1141 GATGCAATGCTGATTTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
 DB 1141 GATGCAATGCTGATTTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
 QY 1201 GAAATCCGATTTATTTTATATATGTTTATCTGTTGTTGTTGTTGTTGTTGTTGTTG 1260
 DB 1201 GAAATCCGATTTATTTTATATATGTTTATCTGTTGTTGTTGTTGTTGTTGTTGTTG 1260
 QY 1261 GTACTTGAAGAAACCAAGAACTATATATATTA 1299
 DB 1261 GTACTTGAAGAAACCAAGAACTATATATATTA 1299

RESULT 5

AAD02788/c
 ID AAD02788 standard; DNA; 577 BP.

XX AAD02788;
 AC 31-MAY-2001 (first entry)
 DT 31-MAY-2001 (first entry)
 DE S. cerevisiae ERG8 gene homologous DNA #2 from *Candida albicans*.
 XX Phosphomevalonate kinase; PKM; ERG8; anti-fungal agent; diagnosis;
 KW infection; ds.
 OS *Candida albicans*.
 XX WO200114533-A2.
 PN 01-MAR-2001.
 PD 15-AUG-2000; 2000WC-GB003100.
 PF 21-AUG-1999; 99GB-00019766.
 PR (ASTR) ASTRAZENECA AB.
 PA (ASTR) ASTRAZENECA UK LTD.
 XX Rosamond JDC, Schnell NF;
 XX WPI; 2001-218441/22.
 DR New polypeptides and polynucleotides (ERG8) from *Candida albicans*, useful
 PT in assays for identifying inhibitors of phosphomevalonate kinase activity
 PT and as reagents for diagnosing *C. albicans* infection.
 XX Claim 6; Page 25; 29pp; English.
 XX

CC The patent discloses phosphomevalonate kinase (PMK; ERG8) protein and
 CC their corresponding DNAs from *Candida albicans*. The ERG8 protein is
 CC useful in assays for identifying compounds that inhibit phospho-
 CC mevalonate kinase (PMK) activity. These inhibitors are useful as anti-
 CC fungal agents. The ERG8 DNA and protein are also useful as reagents for
 CC diagnosing *C. albicans* infection. The present sequence is a *Candida*
 CC *albicans* DNA which is homologous to the *Saccharomyces cerevisiae* ERG8
 CC gene

XX Sequence 577 BP; 167 A; 109 C; 108 G; 190 T; 0 U; 3 Other;

Query Match 40.1%; Score 521; DB 4; Length 577;
 Best Local Similarity 97.1%; Pred. No. 2,6e-129;

Matches 561; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 177 AAACGAGAAATGGGAATATCATATCATCAAAATACAGAACCCAGAGAGATTGACG 236
 DB 577 AACGGGGAATGGGAATATCATATCATCAAAATACAGAACCCAGAGAGATTGACG 519
 QY 237 ACCGATTAAT-CCATTTTAAAGGCAACTATATTCATCG--TTTAACTTAATTCAC 293
 DB 518 ACCGATTAATCCATTTTAAAGGCAACTATATTCATCGGTTTAACTTAATTCAC 459
 QY 294 GACCGAAGCATTTGATCTGAATCATCATCTTCTGAGACCCCTGGATTCATTCAACA 353
 DB 458 GACCGAAGCATTTGATCTGAATCATCATCTTCTGAGACCCCTGGATTCATTCAACA 399
 QY 354 AGATACTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTACCATTCCTGTC 413
 DB 398 AGATACTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTACCATTCCTGTC 339
 QY 414 CATTACCGAAGTGGAAAAACCGGATTAAGTTGATCGCAGAGATTAGTCTCAGTTGTC 473
 DB 338 CATTACCGAAGTGGAAAAACCGGATTAAGTTGATCGCAGAGATTAGTCTCAGTTGTC 279
 QY 474 CACAAATTATTAATCCATTTTATCCCAATGTTATCAGTAAGAAATTAATTTGCA 533
 DB 278 CACAAATTATTAATCCATTTTATCCCAATGTTATCAGTAAGAAATTAATTTGCA 219
 QY 534 CAACGTTGACAGATTGCACTTTGTTATGCCAAAAAAAGATAGATCGGTTGATGT 593
 DB 218 CAACGTTGACAGATTGCACTTTGTTATGCCAAAAAAAGATAGATCGGTTGATGT 159
 QY 594 TCCAACTGCAATTTATGATCTGATGTAATAGAAAGTTCCGACAGCTTTGATTAATGA 653
 DB 158 TCCAACTGCAATTTATGATCTGATGTAATAGAAAGTTCCGACAGCTTTGATTAATGA 99
 QY 654 CGTGTTCAGGTTCTAGAAAGTATCTGAGAAAGTTCCGACAGCTTTGAAATAATGAT 713
 DB 98 CGTGTTCAGGTTCTAGAAAGTATCTGAGAAAGTTCCGACAGCTTTGAAATAATGAT 39
 QY 714 TGAAGTAACTGGGAATTCAAACATGAAAGATGTACAT 751
 DB 38 TGAAGTAACTGGGAATTCAAACATGAAAGATGTACAT 1

RESULT 6
 AAD02787 standard; DNA; 547 BP.

ID AAD02787
 XX AAD02787;

DT 31-MAY-2001 (first entry)

DE S. cerevisiae ERG8 gene homologous DNA #1 from *Candida albicans*.

KW Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;

XX Candida albicans.

XX OS
 XX PN
 XX MO000114533-A2.

PD 01-MAR-2001.

XX 15-AUG-2000; 2000WO-GB003100.

XX 21-AUG-1999; 99GB-00019766.

XX (ASTR) ASTRAZENECA AB.

XX (ASTR) ASTRAZENECA UK LTD.

XX Rosamond JDC, Schnell NF;

XX WPI; 2001-218441/22.

PT New polypeptides and polynucleotides (ERG8) from *Candida albicans*, useful
 PT in assays for identifying inhibitors of phosphomevalonate kinase activity
 PT and as reagents for diagnosing *C. albicans* infection.

XX Claim 6; Page 24; 29pp; English.

CC The patent discloses phosphomevalonate kinase (PMK; ERG8) protein and
 CC their corresponding DNAs from *Candida albicans*. The ERG8 protein is
 CC useful in assays for identifying compounds that inhibit phospho-
 CC mevalonate kinase (PMK) activity. These inhibitors are useful as anti-
 CC fungal agents. The ERG8 DNA and protein are also useful as reagents for
 CC diagnosing *C. albicans* infection. The present sequence is a *Candida*
 CC *albicans* DNA which is homologous to the *Saccharomyces cerevisiae* ERG8
 CC gene

Sequence 547 BP; 184 A; 81 C; 123 G; 159 T; 0 U; 0 Other;

Query Match 36.2%; Score 470.4; DB 4; Length 547;
 Best Local Similarity 99.8%; Pred. No. 9.5e-116;

Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 828 CCAATGAAAAAGGAAAAAGCCAGAAAGCTCTGTTGATGATGACCACTTAATGTGC 887
 DB 1 CCAATGAAAAAGGAAAAAGCCAGAAAGCTCTGTTGATGATGACCACTTAATGTGC 60
 QY 888 CAATTTACGTTTATGAAGGATTTGAGGGAATGCGTGAATAATAGCACTCAACCCAGA 947
 DB 61 CAATTTACGTTTATGAAGGATTTGAGGGAATGCGTGAATAATAGCACTCAACCCAGA 120
 QY 948 GACTTATATTAAAGAGTTAGATCATTCGTTGAGCCTTGACTGTTGCAATTAAAGACAT 1007
 DB 121 GACTTATATTAAAGAGTTAGATCATTCGTTGAGCCTTGACTGTTGCAATTAAAGACAT 180
 QY 1008 CAGAAAAAGGTTTCAAGCATTAACAAAAATAGAGTTTCAATTGAACCTGATGTCCA 1067
 DB 181 CAGAAAAAGGTTTCAAGCATTAACAAAAATAGAGTTTCAATTGAACCTGATGTCCA 240
 QY 1068 AACCCAGTTGTTGACCGTTGCAAGAGATTCCTGCTGTTGTTGGTGGTGTCCAGG 1127
 DB 241 AACCCAGTTGTTGACCGTTGCAAGAGATTCCTGCTGTTGTTGGTGGTGTCCAGG 300
 QY 1128 TGCTGTGATACGATGCAATAGCTGTATTAAGTTGAAAAATCAAGTGGAAAAATTTTAA 1187
 DB 301 TGCTGTGATACGATGCAATAGCTGTATTAAGTTGAAAAATCAAGTGGAAAAATTTTAA 360
 QY 1188 GCAGAAAAGCTTTGAAAAATCCGATTAATTTTCAATAAGTTTCTGGGTTGATTTGGAGCA 1247
 DB 361 GCAGAAAAGCTTTGAAAAATCCGATTAATTTTCAATAAGTTTCTGGGTTGATTTGGAGCA 420
 QY 1248 GCAGAACAGAGGTGTACTTGAAGAAAAACAGAGACTATATAGGTTTATAA 1299
 DB 421 GCAGAACAGAGGTGTACTTGAAGAAAAACAGAGACTATATAGGTTTATAA 472

RESULT 7

AAL40805 standard; DNA; 1356 BP.

XX AAL40805;
 XX AC

DT	03-OCT-2002	(first entry)
DE	Nucleic acid relating to the production of prenyl alcohol SEQ ID No 29.	
XX		
XX		
XX	Prenyl alcohol; mutated cell; squalene synthase gene; geometric isomer;	
KM	industrial synthesis; isoprenoid-terpenoid compound; db.	
XX		
OS	Saccharomyces cerevisiae.	
XX		
PN	WO200253747-A1.	
XX		
PD	11-Jul-2002.	
XX		
XX	20-DEC-2001; 2001WO-JP011215.	
PF		
XX	28-DEC-2000; 2000JP-00401701.	
PR	28-DEC-2000; 2000JP-00403067.	
PR	18-SEP-2001; 2001JP-00282978.	
XX		
PA	(TOYT) TOYOTA JIDOSHA KK.	
XX		
PI	Ohno C, Obata S;	
XX		
DR	WPI, 2002-548086/58.	
XX		
PT	Production of prenyl alcohols by culturing translationally-active mutated	
PT	cells with reduced squalene synthase gene to express less transcriptional	
PT	product.	
XX		
PS	Disclosure; Page 237-238; 262pp; Japanese.	
XX		
CC	The invention relates to a method for producing prenyl alcohol comprising	
CC	culturing mutated cells, having been mutated so as to reduce the amount	
CC	of transcriptional product of squalene synthase gene transcriptional	
CC	activity, and then collecting prenyl alcohol from the culture medium. The	
CC	method is for the production of prenyl alcohols, which is for use in	
CC	industrial synthesis of isoprenoid-terpenoid compounds particularly	
CC	physiologically-active prenyl alcohol geometric isomers. This	
CC	polynucleotide sequence represents a nucleic acid sequence relating to	
CC	the method for producing prenyl alcohol comprising culturing mutated	
XX	cells	
XX		
XX	Sequence 1356 BP; 427 A; 236 C; 304 G; 389 T; 0 U; 0 Other;	
SO		
Query Match	10.2%; Score 132; DB 6; Length 1356;	
Best Local Similarity	50.5%; Pred. No. 8, 1e-25;	
Matches	473; Conservative 0; Mismatches 430; Indels 33; Gaps 5	
QY	3 GTCAAAAGCATTAGTGCACCTGGAAGAGCATTTCTTGCTGTGATATTGGTCTTGA 62	
DB	9 GTTGAGACCTTCAGTGCAGCCCGGAGGAAAGCTTACTAGCTGGTGATATTAGTTTAA 68	
QY	63 GCCATTTATGATGCTTATGTGTACAGCATTTGTCATACGAATGCACTGTTATACAC 122	
DB	69 TACAAAATATGAAAGCATTTGTGTGCGGATTTTCGCAAGAAATGCATGCTAGCCATCC 128	
QY	123 AAAAGAAC-----CAGTTGAAAGAACTGAAATCAAAATTTCTCACCCCAATTGC 176	
DB	129 TTACGCTTCATTGCAAGGCTCTGATATAGTTTGAAGTCGCTGCAAAAGTAAACATTTAA 188	
QY	177 AAACGGAAGATGGAAATATCACATATCATCAATACAG--AGAGCCGAGAGAAATTCA 233	
DB	189 AGATGGGAGTGGCTGTATCATATATAGCTCTAAAGATGCTTCATTCTGTTTCGATAG 248	
QY	234 GTACAGCATTAATCCATTTTATAGAGCAACTATATTCATGCTTTATGCTTATTTCAAC 293	
DB	249 CGGATCTTAAGAACCTTTTCATTGAAAAGATTATCGCTAACGTATTTAGCTTAAAC 308	
QY	294 GACCGAAGCATTTGATCTTGAAG-----TCATCATTTTACTCAGACCTGATATCA 344	
DB	309 TAACTATGACGACTACTGCAATATGAAACTTTGCTTATATGATATTTTCTCTGATGATGC 368	
QY	345 TTGACAAAGATGATCTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTTACA 404	

Dd		369 CTACACTTCTCAGGAGATAGCGTTAACCGAACAATGCTGGCAACGAAGAATTGAGTTTCA 428
Oy		405 TTCTGTGCCTACCTTACCAGAGTGAAAAAGACCCGATTTAGGTTCA TCGGACGAAATTAGTTC 464
Dd		429 TTCGACACGAATTAAGAAAGTTCCCAAACAGGGCTCGGCTCCTCGGACAGGTTTAGTCAC 488
Oy		445 AGTTGTGCCACAGTTTATTAATCCATTTTATCCC-----CAATGTATCACTAC 515
Dd		489 AGTTTAACTACAGCTTGGCCCTCTTTTTTTTATTCGACCTGGAAAATTAATGTAGCAA 548
Oy		516 GAATAAGATATTTTGCAACAAGTTGCAACAGATFGCACATTTGTTATGCCCAAAAAAGAT 575
Dd		549 ATATAGAGAGTTATTCATTAATTTAGCACAGATGCTCATGTCAAGCTCAGGGTAAAT 608
Oy		576 AGGATCTGGGTTTGATGTTGCAACTGCAATTTATGCTCATGTTATATAGAAATTTCA 635
Dd		609 TGGAAGCGGGTTTGATGTATGCGGCGGACAGATATGATCTATCAGATATAGAAAGATTCCC 668
Oy		636 GCCAGCTTTGATAAATGACGTTGTTCAAGTTCTAGAAATGATCTTGAGAATTTCCCCAC 695
Dd		669 ACCCCCACTTAATCTCAATTGCCAGATATTT-----GGAAAGGCTAACCTTACGCGAGTAA 722
Oy		686 AGAGTTGAAAAAATGATTAAGTAAGTAACCTGGGAATTCAAAACATGAAGAATGTACATTACC 755
Dd		723 ACTGGCGCATTTGGTTGATGAAGAAAGACTGGAATATATTACATTAAGTAACCAATTACC 782
Oy		756 ATACGGAATCAAGTATTATTAATGGGTGACGTCACAAAGGTGGCTCAGAAACACCCAATTTGCT 815
Dd		783 TTCGGGATTAACCTTTATGATGATGGCGATATTAAGAATGCTTCAGAAACAGTAABAAGTGT 842
Oy		816 ATCAGCAGTACTCCAATGGAAGAAAGAAAGCCAGAAAGAGCTCTGTGTGTATGACCA 875
Dd		843 CCAGAGGATGAAGAAATTTGGTATGATTCGCATATGCCAGAAAGCTGAAAAATATATACAGA 902
Oy		876 GCTTAATAGTGCCTATTTACAGTTTATGAAGAAATT 911
Dd		903 ACTGCATCATGCAAAATTCTAGATTTATGATGATGACT 938
 RESULT 8 AADJ1013 ID AADJ1013 standard; DNA, 1356 BP.		
XX	AADJ1013;	
DT	31-MAY-2002 (first entry)	
DE	Yeast phosphomevalonate kinase (ERG8) orf.	
KM	Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;	
KW	isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;	
KX	transgenic plant; yeast; phosphomevalonate kinase; ERG8; PMK; ds.	
OS	Saccharomyces cerevisiae.	
PN	WO200210398-A2.	
PD	07-FEB-2002.	
PF	31-JUL-2001; 2001WO-US024037.	
PR	31-JUL-2000; 2000US-0221703P.	
PA	(HAHN/) HAHN F M.	
PA	(KUEH/) KUEHNLE A R.	
PI	Hahn FM, Kuehnle AR;	
DR	WPI, 2002-217122/27.	
Use of specific genes of mevalonate and isoprenoid biosynthetic pathways, for providing a cell with herbicide or antibiotic resistance, and for		

PT providing transformed cells having increased isoprenoid production.
 XX Disclosure; Page 116-117; 193pp; English.
 XX The invention relates to the use of specific genes of the mevalonate and
 CC isoprenoid biosynthetic pathways and inactive gene sites (pseudogenes).
 CC Genes of the invention are used to enhance biosynthesis of isoprenoid
 CC diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid
 CC pathway derived products in the plastids of transgenic plants and
 CC microalgae, for producing herbicide or antibiotic resistant transgenic
 CC plants and microalgae, for providing transformed cells with increased
 CC isoprenoid production compared to non-transformed cells, and for
 CC providing a cell with an inserted polynucleotide sequence encoding one or
 CC more products of interest. The present sequence is yeast
 CC phosphomevalonate kinase (PMK; ERG8) EC 2.7.4.2. encoding orf
 XX
 SQ Sequence 1356 BP; 427 A; 236 C; 304 G; 389 T; 0 U; 0 Other;
 Query Match 10.2%; Score 132; DB 6; Length 1356;
 Best Local Similarity 50.5%; Pred. No. 8.1e-25;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
 QY 3 GTCAAAAGCATTTAGTACCTGGAAAAGCATTTCTTGGTGGGATATTGGTCTTGA 62
 Db 9 GTTGAGAGCCTTACGTGCCCCAGGAAAGCTTACTGCTGGGATATTAGTTTAA 68
 QY 63 GCCAATTTATGATGCTTATGTACAGCATTTGTCATCGAATGATGCAATTAACACC 122
 Db 69 TACAAATATGAAACATTTGTAGTCGATATTCGCAAGAAATCATGCTGTAGCCATCC 128
 QY 123 AAAAGAAC-----CAGTTGAAAGATCTAATCAAAATTTCTTCAACCCCATTTGC 176
 Db 129 TTACGGTTCATTCGCAAGGCTCTGATTAAGTTGAAGTCCTGTGAAAGTAAACATTTAA 188
 QY 177 AAACGGAATGGGAATATCATATCATCAATACAG--AGAAAGCCAGAGAAAGTTCA 233
 Db 189 AGATGGGAGTGGCTGTACATTAAGTCTTAAAGTGGCTTATCTCGTTTCATAG 248
 QY 234 GTACGCAATTAATTCATTTTGAAGCACTATATTCATCGTTTAACTTATATTTCAAC 293
 Db 249 CGGATCTAAGAACCTTCTTCAATGAAAGTTATCGCTAAGCTTATAGCTTTAAAC 308
 QY 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTTACGAGCCCGGATATCA 344
 Db 309 TTAACATGACGACTACTGCAATGAAGAACTGTTCGTATGATATTTCTCGATGATGC 368
 QY 345 TTCACAGAAGATCTGAAACCAAGACATCTCGAATGAGAAAACATTTCTTTACCA 404
 Db 369 CTACCATTTCTAGAGGATAGCGTTACCGAACATCGTGGCAACAGAAATGAGTTTCA 428
 QY 405 TTCTGTCGCAATTAACGAGTGAAGAAAGCCGATTTAGTTCATCGGAGATTAAGTGC 464
 Db 429 TTCGACAGAAATGAAGAAAGTCCCAAAAGAGGCTGAGCTCTCGGAGGTTAGTAC 488
 QY 465 AATTGTGCGCAAAATTTATATCCATTTATCC-----CATGTTATCAGTAC 515
 Db 489 AATTTTAATACAGCTTGGCTCTTTTGTATCGAGCCGMAAATATATGACAA 548
 QY 516 GAATAAGATATTTGGCAACAGTGGCAAGATTGCAATTTGATATGAGAAATGAT 575
 Db 549 ATATAGAGAGATTAATTAATTAATTAAGCAAGTGGCTATTTGTAAGCTGAGGTAAT 608
 QY 576 AGGATCGGCTTATGATGTTGCAACTGCAATTTATGCTGATTTATATGAGAAATTTCA 635
 Db 609 TGGAGCGGGTGTATGATGAGCGGCGGAGCATATGATCTATGATTAAGAAATTTCC 668
 QY 636 GCCAGCTTGAATATGACGCTTTACAGTCTTGAAGAAAGTATCTGAGAAATTTCCAC 695
 Db 669 ACCGCAATTAATCTTAATTTGCGAGATTT-----GAAAGTCTACTTACCGCAGTAA 722
 QY 696 AGAGTTGAAAAAATTTGATGAAAGTAACTGGGAATTCAAACGAAAGATGATACATTACC 755
 Db 723 ACTGGCGCATTTGGTGTATGAAGAACTGGAATATTAATTAATTAAGTAAACATTTACC 782

QY 756 ATACGGAATCAAGTTATTAATGATGAGCTGACAGGAGTGTCTGAGAAACCAAAATTGCT 815
 Db 783 TTCGGGATTAATCTTATATGATGAGGAGCATATTAAGAAATGTTCAAGAAACGTAATCTGCT 842
 QY 816 ATCAGAGATGATCTCCATGAAAGAAAGCCAGAAAGCTCTGTTGTATGACCA 875
 Db 843 CCAAGAGGTAAAGAAATTTGTATGATTCGCAATATGCGCAAGAAAGCTTGAAGAAATATATACGA 902
 QY 876 GCTTAATAGTCCATTTACAGTTTATGAGAAATTT 911
 Db 903 ACTGCATCATGCAATTTCAATTTATGATGAGTACT 938
 RESULT 9
 ABR96801
 ID ABR96801 standard; DNA; 1356 BP.
 AC ABR96801;
 XX
 XX 24-SEP-2002 (first entry)
 DT
 XX
 DE 5. cerevisiae prenyl diphosphate synthase gene #20.
 KW Prenyl alcohol; prenyl diphosphate synthase; geranylgeraniol;
 KW hydroxymethylglutaryl-coA reductase; farnesylgeraniol; gene; ss;
 KW isopentenyl diphosphate delta-isomerase; mevalonate kinase;
 KW mevalonate CoA acetyltransferase; isoprenoid-terpenoid compound.
 XX
 OS Saccharomyces cerevisiae.
 PN
 PN WO200253746-A1.
 PD 11-JUL-2002.
 PD 20-DEC-2001; 2001WO-JP011214.
 PF 28-DEC-2000; 2000JP-00403067.
 PR
 XX
 XX (TOYT) TOYOTA JIDOSHA KK.
 PA
 PI Ohto C, Obata S, Muramatsu M, Nishi K, Totusuka K;
 DR WPI; 2002-537944/57.
 DR
 XX
 PT Production of prenyl alcohols by culturing a transformant transferred
 PT with e.g. prenyl diphosphate synthase gene, for use in industrial
 PT synthesis of e.g. physiologically-active isoprenoid-terpenoid compounds.
 PS
 PS Example 17; Page 269-270; 335pp; Japanese.
 XX
 CC The invention relates to a process for producing a prenyl alcohol
 CC comprising: (A) construction of a recombinant by transferring an
 CC expression recombinant DNA or a DNA for genome integration into a host
 CC which contains prenyl diphosphate synthase gene or its variant; and (B)
 CC collecting product from the culture medium. Also described is: (1) a
 CC method for producing prenyl alcohol in which the expression recombinant
 CC DNA contains: (a) a hydroxymethylglutaryl-CoA reductase gene or its
 CC variant; or (b) an isopentenyl diphosphate delta-isomerase gene; (2) a
 CC process for producing geranylgeraniol, comprising: (a) constructing a
 CC recombinant by transferring an expression recombinant DNA or a DNA for
 CC genome integration into a host which contains hydroxymethylglutaryl-CoA
 CC reductase gene or its variant; and (b) isolating the product; (3) a
 CC process for producing farnesylgeraniol in which the expression
 CC recombinant DNA also contains an isopentenyl diphosphate delta-isomerase
 CC gene; mevalonate CoA acetyltransferase gene, hydroxymethylglutaryl-CoA
 CC synthase gene, mevalonate kinase gene, or mevalonate diphosphate
 CC decarboxylase gene. The methods are used for the production of prenyl
 CC alcohols, particularly for use in industrial synthesis of isoprenoid-
 CC terpenoid compounds e.g. physiologically-active prenyl alcohols including
 CC geometric isomers. ABR96780-ABR96897 represent prenyl diphosphate
 CC synthase genes and related PCR primers used in the methods of the
 CC invention

XX Sequence 1356 BP; 427 A; 236 C; 304 G; 389 T; 0 U; 0 Other;

Query Match 10.2%; Score 132; DB 6; Length 1356;
Best Local Similarity 50.5%; Pred. No. 8.1e-25;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

```

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAGAACATTTCTTGCTGGTGTGATATTTGGTCTTGA 62
DB 9 GTTGAGAGCTTCAAGTGGCCCAAGGAAAGCGTTACTAGCTGGTGGATATTTAGTTTAA 68
QY 63 GCCAATTATGATGCTTATGACAGCATTTGTCATCAAGATGATGACAGTTATTAACCC 122
DB 69 TACAAATATGAAGCATTTGTGTCGATTAATCGGCAAGAAATGCAATGCTGTAGCCCATCC 128
QY 123 AAAAGAAC-----CAGTTGAAGAACTAGAAATCAAAATTTCTTCAACCCCAATTTC 176
DB 129 TTACGGTTCATTGCAAGGGTCTGATTAAGTTGAAGTGGCTGTGAAAGTAAACAAATTTAA 188
QY 177 AAACGGAGATGGGAATATCATATATCATCAATAACAG---AGAAAGCCAGAGAAAGTTCA 233
DB 189 AGATGGGAGTGGCTGTACATATTAAGTCTTAAGAGTGGCTTCATTCCTGTTCCATAG 248
QY 224 GTCACGCATTAATCCATTTTAAAGGCAACTATATTCATGCTTTAGCTTAATTCACCC 293
DB 249 CGGATCTAAGAACCTTTTCATTGAAAGAAAGTATCGCTAAGTATTAAGCTTAAAC 308
QY 294 GACCGAGATTTGATCTTGAA-----TCATCATTTACTGACAGCCCTGATATCA 344
DB 309 TAAACATGACGACTCTGCAATTAAGAACTGTGCTGATTTGATTTCTGATGATGC 368
QY 345 TTCACAAAGAACTAGTAAACCAAGCATCTCGAATGAGAAAGAAACATTTCTTTACCA 404
DB 369 CTACCATTTCTCAGAGAGATAGGCTTACCAACATGCTGGCAACAGAAATTTAGTTTCA 428
QY 405 TTCTGTCGCAATTAACGAAAGTGAAGAAAGCCGATTAAGTTGATTCGCGACAGATTAGTGC 464
DB 429 TTCGACAGAAATTAAGAAAGTTCCCAAAACAGGGCTGGCTCTCGGACGTTTATGCTAC 488
QY 455 AGTTGTGCGACAAAGTTTATTCATTTTATCC-----CAATGTTATCACTAC 515
DB 489 AGTTTAACTACAGCTTGCGCTCTCTTTTGTATCGACCTGGAAGAAATATGTAGACA 548
QY 516 GAATTAAGATATTTTGCACACGTTGCACAGATTTGCAATGTTATGCCCCAAAAAGAT 575
DB 549 AATATGAGAGATTTATTCATTAATTTAGCACAAGTTGCTCATTTGCAAGCTCAGGTAAT 608
QY 576 AGGATCTGGGTTGATGTTGCAACTGCAATTTATGCTGATTTGATATAGAAATTTCA 635
DB 609 TGGAGCGGGTTGATGTAAGCGCGGACACATATGATCTATCAATATAGAAATTTCC 668
QY 636 GCCAGCTTGAATTAATGACGTGTTTCAAGATTTCTGAAAGTATCCTGAGAAAGTTCCCC 695
DB 669 ACCCCCATTAATCTCTAATTTGCCAGATATTT-----GGAAGGCTACTTACGGCAGTA 722
QY 696 AGAGTTGAAGAAATTTGATTAAGTAAGTAACTGGAATTCACAAACATGAAGATTAATACC 755
DB 723 ACTGGGCACTTTGTTGATTAAGAAAGACTGGAATTTAGCATTTAAAGTAACCATTTACC 782
QY 756 ATACGGAATCAAGTATTAATTAAGGTCGACCTCAAGGGTGGCTGAGAAACCCCAATTTGT 815
DB 783 TTGGGATTAATCTTATATGATGAGGCGATTTAAGATGTTAGAAACGTAATGAT 842
QY 816 ATCAGAGTATCTCAATGGAAGAAAGAAAGCCAGAAAGAGCTCTGTTGTGATGACCA 875
DB 843 CCAGAAATATGAAGAAATTTGATGATTCGATATGCCCAAGAGCTGGAAGAAATATATACAG 902
QY 876 GCTTAATAGGCAATTTACAGTTATGAAGGAAT 911
DB 903 ACTCGATCATGCAAAATTTAGATTTATGATGACT 938

```

RESULT 10

```

ADP23287
ID ADP23287 standard; DNA; 1356 BP.
XX
AC ADP23287;
XX
DT 12-FEB-2004 (first entry)
XX
DE MevB operon phosphomevalonate kinase (PMK) gene.
XX
KW isopentenyl pyrophosphate; mevalonate pathway;
KW isopentenyl pyrophosphate synthesis; isoprenoid; pharmaceutical;
KW nucleotides; flavouring agent; agricultural pest control agent;
KW acetoacetyl-CoA thiolase; atob; mevalonate bottom operon; MevB operon;
KW gene; ds; phosphomevalonate kinase; PMK.
XX
OS Bacteria.
XX
PN US2003148479-A1.
XX
PD 07-AUG-2003.
XX
PF 06-DEC-2001; 2001US-00006909.
XX
PR 06-DEC-2001; 2001US-00006909.
XX
PA (KEAS/) KEASLING J.
PA (MART/) MARTIN V.
PA (PITE/) PITEIRA D.
PA (KIMS/) KIM S.
PA (WITH/) WITHERS S T.
PA (YOSH/) YOSHIKUNI Y.
PA (NEWM/) NEWMAN J.
PA (KHLE/) KHEBNIKOV A V.
XX
PI Keasling J, Martin V, Pitera D, Kim S, Withers ST, Yoshikuni Y;
PI Newman J, Khebnikov AV;
XX
DR WPI; 2004-020454/02.
XX
PT Synthesizing isopentenyl pyrophosphate in a host microorganism, useful
PT for pharmaceutical purposes, comprises introducing into the microorganism
PT heterologous nucleic acid sequences coding for an enzyme in the
PT mevalonate pathway.
XX
PS Claim 11; SEQ ID NO 5; 40pp; English.
XX
CC The invention describes a method of synthesizing isopentenyl
CC pyrophosphate (I) in a host microorganism. The method comprises
CC introducing into the host microorganism a plurality of heterologous
CC nucleic acid sequences each coding for a different enzyme in the
CC mevalonate pathway for producing isopentenyl pyrophosphate. The method is
CC useful in synthesizing isopentenyl pyrophosphate and the derived
CC isoprenoids. The isoprenoids may be used in pharmaceuticals,
CC nutraceuticals, flavouring agents or in agricultural pest control agents.
CC This sequence represents the phosphomevalonate kinase (PMK) gene from the
CC Mevalonate bottom (MevB) operon. Note: The specification states that MevB
CC and MevT operon gene are isolated from Saccharomyces cerevisiae and
CC Escherichia coli but does not state which species this polynucleotide was
CC isolated from.
XX
SQ Sequence 1356 BP; 426 A; 236 C; 305 G; 389 T; 0 U; 0 Other;

```

Query Match 10.2%; Score 132; DB 12; Length 1356;
Best Local Similarity 50.5%; Pred. No. 8.1e-25;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

```

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAGAACATTTCTTGCTGGTGTGATATTTGGTCTTGA 62
DB 9 GTTGAGAGCTTCAAGTGGCCCAAGGAAAGCGTTACTAGCTGGTGGATATTTAGTTTAA 68
QY 63 GCCAATTATGATGCTTATGACAGCATTTGTCATCAAGATGATGACAGTTATTAACCC 122
DB 69 TACAAATATGAAGCATTTGTGTCGATTAATCGGCAAGAAATGCAATGCTGTAGCCCATCC 128

```

```

QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGAAATCAAAATTTCTTCACCCCAATTGG 176
DB 129 TTACGGTTTCATGCAAGGGTCTGATAGTTTGAAGTCCGTGTAAGAAAGTAAACATTTAA 188
QY 177 AAACGGAATGGGAATATCATATCATCAATATACG---AGAACCCAGAGAGTTCA 233
DB 189 AGATGGGGAGTGGCTGTACATATTAAGTCTAAAGGGCTTCATTCGTGTTGATAGG 248
QY 234 GTACGCAATTAATCCATTTTAAAGGCACTATATTCATGCTTTAGCTTAATTTCAACC 293
DB 249 CGGATCTAAGAACCTTTCTTCAAAAAAGTTATGCTTAACGATTTAGCTACTTTAAACC 308
QY 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTACTCAGACCTTGATATCA 344
DB 309 TTAACATGACGACTACGCAATAGAAAATTGTTGGTATTTGATTAATTTCTGATGATGC 368
QY 345 TTCACAAGAAAGATCTGAAACCAAGACATCTCGAATGAGAAAAACATTTCTTTACA 404
DB 369 CTACCAATTTCTCAGAGAGATAGCGTTACCGAACATCTGTGCAACAGAAATTTGAGTTTCA 428
QY 405 TTCTCGTCATTAACCGAAGTGGAAAAAGACCGGATTAGGTTCAATGGCAGATTAGTGC 464
DB 429 TTTCGACAGATTTGAAGAGTTCCCAAAACAGGGCTGGCTCCTGGCAGGTTTATGTCAC 488
QY 465 AGTTGTTGCCACAAGTTTATTTATCCATTTTATCCC-----CATGTTATCAGTAC 515
DB 489 AGTTTATCTACAGCTTTGGCTCTTTTGTGATGAGACCTGGAATAATATGATGACAA 548
QY 516 GAATAAGATATTTTGCACACGCTTGACAGATGTCACATTTGATGCCCCAAAAAGAT 575
DB 549 ATATAGAGAAATTAATCAATTAATTTAGCACAAGTGTCTATGTTCAAGCTCAGAGTAAAT 608
QY 576 AGGATCTGGGTTTGAATGTTGCAACTGCAATTTATGCTGATTTGATTAAGAAATTTCA 635
DB 609 TGGAAAGCGGTTTGAATGAGCGCGGCGGACATATGATATTAACATATTAAGAAATTTCC 668
QY 636 GCCAGCTTTGATTAATGACGTTGTTCAAGTTCTAAGAAAGTATCTGAGAGTTCCCAAC 695
DB 669 ACCCGCATTAATCTTAATTTGCCAGATATTT-----GGAAATGCTACTTACGGCAGTAA 722
QY 696 AGAGTTGAAAAAATGATTAAGAAAGTAACTGGAATTTCAAACTGAAAGATGACATTTAC 755
DB 723 ACTGGCGCTTTGGTTGATGAGAAAGACGTGAATATTAAGATTTAAAGTAACCATTTACC 782
QY 756 ATACGGAATCAAGTATTAATGAGTGAAGCTCAAGGTTGGCTCAGAAACCCCAATTTGAT 815
DB 783 TTCCGGATTAATCTTATGATGGGCGATTTTAAGAAATGTTTCAAGAAACGTTAAACTGCT 842
QY 816 ATCAGAGTACTTCAATGGAAGAAAGAAAGCCAGAAAGAGCTCTGTTGTATGACCA 875
DB 843 CCAGAAAGTAAAAATTTGATATGATTCGATATGCGCAGAAAGCTTGAAAAATATATACGA 902
QY 876 GCTTAATAGTCCCAATTTTACAGTTTATGAGGAAT 911
DB 903 ACTGATCATGCAATTTCAAGTTATATGATGACT 938

```

RESULT 11

AD182315 standard; DNA; 1356 BP.

AD182315;

22-APR-2004 (first entry)

S cerevisiae phosphomevalonate kinase (PMK) gene.

amorph-4,11-diene synthase; mevalonate pathway;

isopentenyl pyrophosphate; amorph-4,11-diene synthase gene;

pharmaceutical; nutraceutical; flavouring agent;

agricultural pest control agent; epi-cedrol; phosphomevalonate kinase;

PMK; gene; de.

```

XX Saccharomyces cerevisiae.
OS
XX US2004005678-A1.
EN
XX 08-JAN-2004.
PD
XX 09-APR-2003; 2003US-00411066.
PF
XX 06-DEC-2001; 2001US-00006909.
PR
XX (KEAS/) KEASLING J.
PA (MART/) MARTIN V.
PA (PIRE/) PITERA D.
PA (WIT/) WITHERS S T.
PA (NEWM/) NEWMAN J.
PI Keasling J, Martin V, Pitera D, Withers ST, Newman J;
PI WPI; 2004-120864/12.
DR
XX
XX
XX This invention relates to a novel method of synthesizing amorph-4,11-
XX diene in a host microorganism which comprises introducing into the host
XX microorganism heterologous nucleic acid sequences, each coding for a
XX different enzyme in the mevalonate pathway for producing isopentenyl
XX pyrophosphate, and introducing into the host microorganism a DNA fragment
XX coding for an optimised synthetic amorph-4,11-diene synthase gene. The
XX nucleic acid sequences, DNA fragments, vectors and host cells are useful
XX for synthesizing isopentenyl pyrophosphate, amorph-4,11-diene and epi-
XX cedrol used as pharmaceuticals, nutraceuticals, flavouring agents, or
XX agricultural pest control agents. The present sequence is that of a gene,
XX encoding an enzyme, which was used during the exemplification of the
XX invention to assemble an operon for the synthesis of amorph-4,11-diene
XX in a host microorganism in the method of the invention.
SQ
Sequence 1356 BP; 426 A; 236 C; 305 G; 389 T; 0 U; 0 Other;
Query Match 10.2%; Score 132; DB 12; Length 1356;
Best Local Similarity 50.5%; Pred. No. 8.1e-25;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
QY 3 GTCAAAAACATTTAGTGCACCTGGAAGAAAGCATTTCTTGCTGTGATATTTGGTTCTTGA 62
DB 9 GTTGAGAGCCTTACGTGCCCCAGGAAAGCGTTACTAGCTGTGAGATTTAGTTTAGA 68
QY 63 GCCAATTTAGATGCTTATGTGACAGCATTTGTCATACGAAATCAGTCAATTAACACC 122
DB 69 TACAAAATATGAAGCATTTGTAGTCGATTAATGCGAAGAAATCAGTGTACCCATCC 128
QY 63 GTCAAAAACATTTAGTGCACCTGGAAGAAAGCATTTCTTGCTGTGATATTTGGTTCTTGA 62
DB 9 GTTGAGAGCCTTACGTGCCCCAGGAAAGCGTTACTAGCTGTGAGATTTAGTTTAGA 68
QY 63 GCCAATTTAGATGCTTATGTGACAGCATTTGTCATACGAAATCAGTCAATTAACACC 122
DB 69 TACAAAATATGAAGCATTTGTAGTCGATTAATGCGAAGAAATCAGTGTACCCATCC 128
QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGAAATCAAAATTTCTTCACCCCAATTGG 176
DB 129 TTACGGTTTCATGCAAGGGTCTGATAGTTTGAAGTCCGTGTAAGAAAGTAAACATTTAA 188
QY 177 AAACGGAATGGGAATATCATATCATCAATATACG---AGAACCCAGAGAGTTCA 233
DB 189 AGATGGGGAGTGGCTGTACATATTAAGTCTAAAGGGCTTCATTCGTGTTGATAGG 248
QY 234 GTACGCAATTAATCCATTTTAAAGGCACTATATTCATGCTTTAGCTTAATTTCAACC 293
DB 249 CGGATCTAAGAACCTTTCTTCAAAAAAGTTATGCTTAACGATTTAGCTACTTTAAACC 308
QY 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTACTCAGACCTTGATATCA 344
DB 309 TTAACATGACGACTACGCAATAGAAAATTGTTGGTATTTGATTAATTTCTGATGATGC 368
QY 345 TTCACAAGAAAGATCTGAAACCAAGACATCTCGAATGAGAAAAACATTTCTTTACA 404

```

```

Db      369 CTACATTTCTCAGAGATAGCGTTATCCGACATCTGTGCAACAGAAAGTTGAGTTTCA 428
Qy      405 TTCTGTCGTCATTAACGAGTGAAGAAAGCCGGATTAAGTTACGCGAGATTGTTGTC 464
Db      429 TTTCGACAGAAATTAAGAAAGTTCCCAAAACAGGGCTGGGCTCTCGGCGATTAGTAC 488
Qy      465 AGTTGTGCGACAGATTATTAATCCCATTTTATCCC-----CATGTTATCAGTAC 515
Db      489 AGTTTAACTACAGCTTTGGCCTCTTTTGTATGAGACCTGGAATAATATGTAACA 548
Qy      516 GAATTAAGATATTTTGCACACGTTGCAACAGATTGCACTGTTATGCCCAAAAAAGAT 575
Db      549 ATATAGAGAGATTATTCATTAATTTAGCAACAGTTGCTCATGTCAGAGCTCAGGGTAAAT 608
Qy      576 AGGATCTGGGTTGATGTTGCAACGCAATTTATGCTGATTTGATATAGAAGATTCA 635
Db      609 TGGAAAGCGGTTTGATGAGCGCGGACGACATATGATCTTATCAGATTAAGAAAGATTCCC 668
Qy      636 GCCAGCTTTGATTAATGACGTGTTTCAGGTTCTAGAAAGTATCCGAAAGTTCCCGAC 695
Db      669 ACCCGCATTAATCTGTAATTTCCAGATAT-----GGAAGTCTACTTACGGCAGTAA 722
Qy      696 AGAGTTGAAAAAATGATTGAAAGTAACTGGGAATCAACATGAAGAATGTACATTACC 755
Db      723 ACTGCGCGCATTTGTTGATGAAGAAAGACTGGAATATTAAGATTAAAGTAAACCATTTACC 782
Qy      756 ATACGAATCAAGTTATTAATGAGTGACGTCGAAGGGTGGCTGCAAAACCCAAATTTGT 815
Db      783 TTCCGATTAATCTTTATGATGATGGCGATATTAAGAAATGTTCCGAAACGTAATCTGTT 842
Qy      816 ATCAGAGTACTCCATGAGAAAGAAAGCAAGAAAGCTCTGTTGTATGACCA 875
Db      843 CCAGAAAGTAAAGAAATGTTGATGATTCGCAATATGCCAGAAAGTTGAAATATATACAA 902
Qy      876 GCTTAATAGTCCCAATTTACAGTTTATGAGAAAT 911
Db      903 ACTGCATCATGCAAAATTTCTAGATTAATGATGATGACT 938

```

RESULT 12

```

ADL14834
ID      ADL14834 standard; DNA; 1356 BP.
XX      AC
XX      ADL14834;
XX      AC
XX      17-JUN-2004 (first entry)
XX      DE
XX      Yeast phophomevalonate kinase gene ERG8.
XX      KM      Yeast; ds; gene; Isopentenyl diphosphate biosynthetic pathway; IPP;
XX      KM      prenyl alcohol; squalene synthase.
XX      OS      Saccharomyces cerevisiae.
XX      PN      US2004063182-A1.
XX      PD      01-APR-2004.
XX      PF      18-JUN-2003; 2003US-00450941.
XX      PR      28-DEC-2000; 2000JP-00401701.
XX      PR      28-DEC-2000; 2000JP-00406067.
XX      PR      18-SEP-2001; 2001JP-00282978.
XX      PR      20-DEC-2001; 2001WO-JP011215.
XX      PA      (OHIO/) OHIO C.
XX      PA      (OBAT/) OBATA S.
XX      PI      Ohio C, Obata S;
XX      DR      WPI; 2004-294401/27.
XX

```

```

PT      Producing prenyl alcohols by culturing a mutant cell in which the amount
PT      of squalene synthase gene transcript having translational activity is
PT      reduced.
PS      Disclosure; SEQ ID NO 29; 139pp; English.
XX      XX
XX      CC      The invention relates to producing a prenyl alcohol comprising culturing
XX      CC      a mutant cell that has been mutated so that an amount of squalene
XX      CC      synthase gene transcript having translational activity can be reduced,
XX      CC      and recovering the prenyl alcohol from the resultant culture.
XX      CC      Alternatively, producing a prenyl alcohol also comprises culturing a
XX      CC      mutant cell in which the transcription promoter region of its squalene
XX      CC      synthase gene has been replaced with a transcription repression-type
XX      CC      promoter under transcription repression conditions, reducing an amount of
XX      CC      squalene synthase gene transcript having translational activity, and
XX      CC      recovering the prenyl alcohol from the resultant culture. Also included
XX      CC      are a method of producing a prenyl alcohol (comprising preparing a
XX      CC      recombinant by introducing a recombinant DNA for expression or a DNA for
XX      CC      genomic integration each comprising an IPP (isoprenyl diphosphate)
XX      CC      biosynthetic pathway-related enzyme gene into any of the mutant cells
XX      CC      cited above), a mutant cell that has been mutated so that an amount of
XX      CC      squalene synthase gene transcript having translational activity can be
XX      CC      reduced) and a mutant cell in which the transcription promoter region of
XX      CC      its squalene synthase gene has been replaced with a transcription
XX      CC      repression-type promoter so that an amount of transcript from its
XX      CC      squalene synthase gene having translational activity can be reduced. The
XX      CC      transcription repression-type promoter in producing a prenyl alcohol is
XX      CC      GAL1 promoter. The IPP biosynthetic pathway-related enzyme gene is any
XX      CC      one selected from farneyl diphosphate synthase gene, geranylgeranyl
XX      CC      diphosphate synthase gene, hydroxymethylglutaryl-CoA reductase gene,
XX      CC      isopentenyl diphosphate A-isomerase gene, mevalonate kinase gene, acetyl-
XX      CC      CoA acetyltransferase gene, hydroxymethylglutaryl-CoA synthase gene,
XX      CC      phosphomevalonate kinase gene and diphosphomevalonate decarboxylase gene,
XX      CC      or a fusion gene composed of a gene selected from the above genes, or a
XX      CC      gene obtained by introducing an addition, substitution or insertion
XX      CC      mutation into any one of the above genes so that the polypeptide encoded
XX      CC      by the resultant gene contains an endoplasmic reticulum signal. The
XX      CC      methods and compositions of the present invention are useful for
XX      CC      producing prenyl alcohols by reducing the amount of squalene synthase
XX      CC      gene transcript having translational activity. The present sequence
XX      CC      encodes an IPP (isoprenyl diphosphate) biosynthetic pathway-related
XX      CC      enzyme.
XX      SQ      Sequence 1356 BP; 427 A; 236 C; 304 G; 389 T; 0 U; 0 Other;
XX
XX      Query Match      10.2%; Score 132; DB 12; Length 1356;
XX      Best Local Similarity 50.5%; Pred. No. 8,1e-25;
XX      Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

```

```

Qy      3      GTCAAAGCATTTAGTACCTGCAAGAAAGCATTTCTGCTGATGATATTTGGTTCTTGA 62
Db      9      GTTGAGAGCTTCATGATGCCCCGAGGAAAGCGTTACTAGCTGTGATATTTAGTTTGA 68
Qy      63      GCCAATTTATGATGCTTATGTCAGCATTTGTCATCAGCAATGATGCAATTATAACACC 122
Db      69      TACAAATATGAAAGATTTTGTAGTCGATATCGGCAAGAAATGATGCTGATGCCATCC 128
Qy      123     AAAAGAAC-----CAGTTGAAGAATCTAGATCAAAATTTCTTACACCCCAATTTGC 176
Db      129     TTACGTTTCATTGCAAGGGTCTGATTAAGTTGAAGTGGCTGTGAAAGTAAACAAATTAA 188
Qy      177     AAAAGGAGAAATGGAATATACATATCAATCAATACAG---AGAAGCCAGAGAAAGTTCA 233
Db      189     AGATGGGAGATGCGCTGATACATTAATGCTTAAAGTGGCTTATCCCTGTTCCATAGG 248
Qy      234     GTCAGCATTAATTCATTTTATAGGCAACTATATTCATGTTTATGCTTATATTTCAACC 293
Db      249     CGGATCTAAGAACCTTTTCATTTGAAGAAAGTATTCCTAAGCTATTAAGTACTTAAACC 308
Qy      294     GACCGAAGCATTTGATCTTTGAAA-----TCAATCTTATCTAGACACCTCGAATATCA 344
Db      309     TAAATGAGACGACTACTGCAATAGAAACTGTTGTTATGATTAATTTCTCTGATGATGC 368

```


QY 345 TTCAAGAAAGATCTGAAACCAAGACATCTGGAATGAGAGAAAAACATTTCTTACCA 404
 Db 369 CTACCAATTTCTCAGAGAGATAGCTTACCGAACAATCTGTGCAACAGAAATTAAGTTTCA 428
 QY 405 TTCTCGATTCATTAACGAGATGGAAGAAAGCCGATTTAGTTTCATGCGAGATTAAGTTC 464
 Db 429 TTGCGACAGAAATTAAGAAAGTTCCCAAAACAGGGCTGCGCTCTCGGAGGTTTGTGAC 488
 QY 465 AGTTGTGCGACAAAGTTTATTTATCCATTTTATCC-----CAATGTTATCAGTAC 515
 Db 489 AGTTTAACTACAGCTTTTGCCCTCTTTTGTATCGACCTCGAAAAATTAATGACAA 548
 QY 516 GAATTAAGATTTTTCACACAGCTTCACAGTTCACATTTGTTATGCCAAAAAAGAT 575
 Db 549 AATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 608
 QY 576 AGGATCTGGGTTTGAATGTTGCACTGCAATTAATGTTGTTGTTATTAAGAAATTTCA 635
 Db 609 TGAAGAGCGGTTTGAATGTAAGCGCGCAGCATATGATCTTATCAGATTAAGAAATTTCC 668
 QY 636 GCGAGCTTTGAATTAATGACGTTTTCAGGTTCTAGAAAGTGATCTGAGAAAGTTCCCGAC 695
 Db 669 ACCCGCATTAATCTCTAATTTGCGCAGATATT-----GAAAGTCTACTTAACGCGAGTAA 722
 QY 696 AGAGTTGAAAAAATGATTGAAAGAACTGGGAATTCAAACATGAAAGATGACATTACC 755
 Db 723 ACTGGCGATTGTTGTTGATGTAAGAAAGATGAAATTAATTAATTAATTAATTAATTAAT 782
 QY 756 ATAGGAAATCAATTAATTAATGAGTGAAGTCAAGAGGTGCTCAGAAACACCCAAATTTGGT 815
 Db 783 TTGCGGATTAATCTTATTAATGAGTGGCGATTAATTAAGAAATGTTCAAGAAACAGTAAATCTGGT 842
 QY 816 ATCAAGAGTACTCAATGGAAGAAAGAAAGCCAGAAAGAAAGCTCTGTTGTTATGACCA 875
 Db 843 CCAAGAGTAAAAAATTTGTAATGATGTTGCGATATGCGCAAGAAAGCTGAAATTAATTAACAGA 902
 QY 876 GCTTAATAGTCCAAATTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 911
 Db 903 ACTGATCATGCAAAATTTCTAGATTTTATGATGATGACT 938

RESULT 13

ADP23291
 ID ADP23291 standard; DNA; 4482 BP.
 AC ADP23291;

DT 12-FEB-2004 (first entry)
 DE Mevalonate pathway operon MewB (Mevalonate bottom).
 XX

KW isopentenyl pyrophosphate; mevalonate pathway;
 KW isopentenyl pyrophosphate synthesis; isoprenoid; pharmaceutical;
 KW nutraceutical; flavouring agent; agricultural pest control agent;
 KW acetacetyl-CoA thiolase; atob; Mevalonate bottom operon; MewB operon;
 KW gene; de; mevalonate kinase; MK; phosphomevalonate kinase; FMK;
 XX mevalonate pyrophosphate decarboxylase; MPD.
 XX

OS Bacteria.
 XX

PN US2003148479-A1.
 XX

PD 07-AUG-2003.
 XX

PF 06-DEC-2001; 2001US-00006909.
 XX

PR 06-DEC-2001; 2001US-00006909.
 XX

PA (KEAS/) KEASLING J.
 PA (MART/) MARTIN V.
 PA (PITE/) PITERA D.
 PA (KIMS/) KIM S.
 PA (WITH/) WITHERS S T.

PA (YOSH/) YOSHIKUNI Y.
 PA (NEMM/) NEMMAN J.
 PA (KHLE/) KHELEBNIKOV A V.
 XX
 PI Keasling J, Martin V, Pitera D, Kim S, WITHERS ST, Yoshikuni Y,
 PI Neman J, Khelebnikov AV,
 XX
 DR WPI; 2004-020454/02.
 XX

PT Synthesizing isopentenyl pyrophosphate in a host microorganism, useful
 PT for pharmaceutical purposes, comprises introducing into the microorganism
 PT heterologous nucleic acid sequences coding for an enzyme in the
 PT mevalonate pathway.
 XX

PS Claim 55; SEQ ID NO 9; 40bp; English.
 XX

CC The invention describes a method of synthesizing isopentenyl
 CC pyrophosphate (I) in a host microorganism. The method comprises
 CC introducing into the host microorganism a plurality of heterologous
 CC nucleic acid sequences each coding for a different enzyme in the
 CC mevalonate pathway for producing isopentenyl pyrophosphate, the method is
 CC useful in synthesizing isopentenyl pyrophosphate and the derived
 CC isoprenoids. The isoprenoids may be used in pharmaceuticals and
 CC nutraceuticals, flavouring agents or in agricultural pest control agents.
 CC This sequence represents the mevalonate pathway operon Mevalonate bottom
 CC (MewB) which contains genes encoding mevalonate kinase (MK),
 CC phosphomevalonate kinase (PMK) and mevalonate pyrophosphate decarboxylase
 CC (MPD). Note: The specification states that MewB and MewT open gene are
 CC isolated from *Saccharomyces cerevisiae* and *Escherichia coli* but does not
 CC state which species this polynucleotide was isolated from.
 XX

Seq Sequence 4482 BP; 1337 A; 928 C; 972 G; 1245 T; 0 U; 0 Other;
 SO

Query Match 10.2%; Score 132; DB 12; Length 4482;
 Best Local Similarity 50.5%; Pred. No. 1.3e-24;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGAACCTCGAAAGAAAGCATTTCTGCTGTGATTAATTTGTTTGA 62
 Db 1606 GTTGAAGAGCTTTAGTGAACCTCGAAAGAAAGCATTTCTGCTGTGATTAATTTGTTTGA 1665
 QY 63 GCGAATTTATGATGCTTATGTAAGCAAGCATTTGATCAAGATGATGATGATTAATTAACCC 122
 Db 1666 TACAAAATATGAAGCATTTGATGATGATTAATGCGAAGATGATGATTAATTAACCC 1725
 QY 123 AAAAGGAC-----CAGTTGAAGATCTGAATCAAAATTTCTTACCCCAATTTGC 176
 Db 1726 TTACGTTATTTGCAAGGCTGATTAAGTTGAAGTGTGTAAGAAAGTAAACATTTAA 1785
 QY 177 AAAAGGAGATGGAATATCAATATCAATATCAATCAATCAATCAATCAATCAATCAATCA 233
 Db 1786 AGATGGAGGATGCTGTAACATATCAATATCAATCAATCAATCAATCAATCAATCAATCA 1845
 QY 234 GTACAGCAATTAATCCATTTTATGAGCACTATATCAATGTTTATGATTAATTAACCC 293
 Db 1846 CGAATCTAAGAACCTTTTCAATTAAGAAAGTATGCTAATGCTATTTAGCTAATTAACCC 1905
 QY 294 GACCGAAGCATTTGATCTTGA-----TCATCAATTAATCAACCTGATATCA 344
 Db 1906 TAAATGAGAGCATCTACATTAAGAACTGTTGTTATTAATTTTCTGATGATGTC 1965
 QY 345 TTCAAGAAAGATCTGAAACCAAGACATCTCGAATGAGAGAAAAACATTTCTTTACCA 404
 Db 1966 CTACCATTTCTCAGAGATAGCTTTACCGAACAATCGTGCAAGAAAGATTAAGTTTCA 2025
 QY 405 TTCTGTCATTTCCGAGTGAAGAAAGCCGATTAATGTTTATCGGACAGATTAAGTTC 464
 Db 2026 TTGCGACAGATTAAGAAAGTTCCCAAAACAGGCTGCGCTCTCGGAGGTTTATGATC 2085
 QY 465 AGTTGTGCGACAAAGTTTATTAATCCATTTTATCC-----CAATGTTATCAGTAC 515
 Db 2086 AGTTTAACTACAGCTTTTGCCCTCTTTTGTATGAGACCTGAAATTAATTAATTAAGACAA 2145

QY 516 GAATAAGATATTTTGCACACAGTTGACAGATTCAGATTGTATGCCCAAAAAAGAT 575
 DB 2146 ATATAGAGAGATTTATTCATTAATTTAGCAACAGTTGCTCATTTGCAAGCTCAGGGTAAAT 2205
 QY 576 AGGATCTGGGTTGATGTTGCACTGCAATTTATGCTCTGATTTGATTAAGAAGATTCA 635
 DB 2206 TGGAAACGGGTTGATGTAAGCCGCGACACATATGATCTATCAGATTAAGAAGATTCCC 2265
 QY 636 GCCAGCTTGAATAATGACGTTTCAGGTTCTAGAAAGTGAATCCGAGAAGTTCGCCAC 695
 DB 2266 ACCCGCATTAATCTCTAATTTCCAGATTTT-----GGAAGTCTACTTACGGCAGTAA 2319
 QY 696 AGAGTTGAAAAAATGATTTGAAGTAACTGGAATTCACATGAAAGATGTACATTACC 755
 DB 2320 ACTGGCGCATTTGGTTGATGAAAGAGACTGGAATATTAAGATTAAGAACATTATACC 2379
 QY 756 ATACGAATCAAGTTATTTATGAGGTGACCTCAAGGTGCTCAAGAAACCCCAATTGGT 815
 DB 2380 TTCGGGATTTAATCTTATGATGCGGATTTTAAGATGCTTCAAGAAACGTAAACTGGT 2439
 QY 816 ATCAGAGTACTCCAATGGAAGAAAGCAAGCAAGAAAGCTGTTGTGATGACA 875
 DB 2440 CCAGAAAGTAAAAAATGATGATTCGATATCCAGAAAGCTGAAAAATATATACAA 2499
 QY 876 GCTTATAGTCCCAATTTACAGTTTATGAAGAATT 911
 DB 2500 ACTGCATCATGCAATTTCTAGATTTATGATGATGACT 2535

RESULT 14

AD182319
 ID AD182319 standard; DNA; 4482 BP.

XX AD182319;

XX 22-APR-2004 (first entry)

DE Mevalonate bottom (MeVB) operon DNA sequence.

XX amorph-4,11-diene synthase; mevalonate pathway;
 XX isopentenyl pyrophosphate; amorph-4,11-diene synthase gene;
 XX pharmaceutical; nutraceutical; flavouring agent;
 XX agricultural pest control agent; epi-cedrol; Mevalonate bottom operon;
 XX MeVB; ds.

OS Saccharomycetes cerevisiae.
 OS Synthetic.

PN US2004005678-A1.

PD 08-JAN-2004.

PF 09-APR-2003; 2003US-00411066.

PR 06-DEC-2001; 2001US-00006909.

PA (KEAS/) KEASLING J.

PA (MART/) MARTIN V.

PA (PITE/) PITERA D.

PA (WITH/) WITHERS S T.

PA (NEWM/) NEWMAN J.

PI Keasling J, Martin V, Pitera D, Withers ST, Newman J;

XX WPI; 2004-120864/12.

PT Synthesizing amorph-4,11-diene in a host cell, useful as
 PT pharmaceuticals, comprises introducing nucleic acid sequences, each
 PT coding for a different enzyme in the mevalonate pathway for producing
 PT isopentenyl pyrophosphate.
 XX Example 1; SEQ ID NO 9; 75bp; English.

CC This invention relates to a novel method of synthesizing amorph-4,11-
 CC diene in a host microorganism which comprises introducing into the host
 CC microorganism heterologous nucleic acid sequences, each coding for a
 CC different enzyme in the mevalonate pathway for producing isopentenyl
 CC pyrophosphate, and introducing into the host microorganism a DNA fragment
 CC coding for an optimised synthetic amorph-4,11-diene synthase gene. The
 CC nucleic acid sequences, DNA fragments, vectors and host cells are useful
 CC for synthesizing isopentenyl pyrophosphate, amorph-4,11-diene and epi-
 CC cedrol used as pharmaceuticals, nutraceuticals, flavouring agents, or
 CC agricultural pest control agents. The present sequence is that of an
 CC operon which may be used in the method of the invention.

XX Sequence 4482 BP; 1337 A; 928 C; 972 G; 1245 T; 0 U; 0 Other;

SQ Query Match 10.2%; Score 132; DB 12; Length 4482;

Best Local Similarity 50.5%; Pred. No. 1.3e-24;

Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAGAACATTTCTGTGTGTGATATTTGGTCTTGA 62
 DB 1606 GTTGAAGCCTTCAGTGCCTCCGAGGAAAGCGTTACTAGCTGTGATATTTAGTTTGA 1665
 QY 63 GCCAATTTATGATGCTTATGATGACAGATTCATCAAGATGATGATGATTAACACC 122
 DB 1666 TACAAAGATTAAGACATTTGATGCGATTAAGCAAGAAATGATGCTGATCCCATCC 1725
 QY 123 AAAAGAAC-----CAGTTGAAAGAAATCTGAATCAAAATTTCTTCACCCCAATTTCG 176
 DB 1726 TTACGCTTCATTCAGAGGCTCTGATTAAGTTTAAGTGCCTGTAAGAAAGTAAACATTTAA 1785
 QY 177 AAACGAGAAATGGGAATATTCACATATCAATTAACAG---AGAAAGCCAGAAAGTTTCA 233
 DB 1786 AATGGGAGAGTGGCTGATACATATTAAGTCTAAAGTGGCTTCATTCCTGTTTCAATAG 1845
 QY 234 GTACGCAATTAATCCATTTTATGAGCACTATTCATCTTTTAACTTATATTCACAC 293
 DB 1846 CGGATCTAAGAACCTTTCATATGAAAGTATTCCTAAGCTTAAGTACTTTTAAACC 1905
 QY 294 GACCGAAGCATTTGATCTTGAA-----TCATCATTTTCTAGACCTCGATATCA 344
 DB 1906 TTAACATGACGACTCTGCAATTAAGAACTTGTCTGATTAATTTTCTGATGATGCTC 1965
 QY 345 TTCACAAAGAAATGACTGAAACCAAGACATCTCGAATGAGAGAAACATTTCTTTACCA 404
 DB 1966 CTACCATCTCAGAGAGATAGGTTACGGAACATGCTGCAACAGAGATTGATTTTCA 2025
 QY 405 TTCTGTGCTCATTCACGAAGTGAACCCGATTAAGTTAGTTATGCGCATTAAGTGTGC 464
 DB 2026 TTCGCAAGAAATGAAAGATTCACCAAAACAGGGCTGGGCTCTCTGGCAGGTTTATGCTAC 2085
 QY 465 AGTTGTGCGACAAGTTTATTTATCCATTTATGCC-----CAATGTTATCACTAC 515
 DB 2086 AGTTTAACTACAGCTTGGCTCTCTTTTATGCGACTGGAATAATATGTAACAA 2145
 QY 516 GAATAAGATATTTTGCACACGTTGACAGATTCAGATTGTATGCCCAAAAAAGAT 575
 DB 2146 ATATAGAGAGATTTATTCATTAATTTAGCAACAGTTGCTCATTTGCAAGCTCAGGGTAAAT 2205
 QY 576 AGGATCTGGGTTGATGTTGCACTGCAATTTATGCTCTGATTTGATTAAGAAGATTCA 635
 DB 2206 TGGAAACGGGTTGATGTAAGCCGCGACACATATGATCTATCAGATTAAGAAGATTCCC 2265
 QY 636 GCCAGCTTGAATAATGACGTTTCAGGTTCTAGAAAGTGAATCCGAGAAGTTCGCCAC 695
 DB 2266 ACCCGCATTAATCTCTAATTTCCAGATTTT-----GGAAGTCTACTTACGGCAGTAA 2319
 QY 696 AGAGTTGAAAAAATGATTTGAAGTAACTGGAATTCACATGAAAGATGTACATTACC 755
 DB 2320 ACTGGCGCATTTGGTTGATGAAAGAGACTGGAATATTAAGATTAAGAACATTATACC 2379
 QY 756 ATACGAATCAAGTTATTTATGAGGTGACCTCAAGGTGCTCAAGAAACCCCAATTGGT 815
 DB 2380 TTCGGGATTTAATCTTATGATGCGGATTTTAAGATGCTTCAAGAAACGTAAACTGGT 2439

Qy	816	ATCCAGAGTACTCCAAATGAGAAAAAGAAAAGCCAGAAAGAGCTCTGTTGTGATGACCA	875
Db	2440	CGAAGAGTAAAAAAATTGATGATTCGATATGCGAGAAAGCTTGAAAAATATACGA	2495
Qy	876	GCTTAATAGTGCCAATTTCACGTTATGAGGAATT	911
Db	2500	ACTGCATCATGCAAAATTCGTGATTTATGATGACT	2535
RESULT 15			
ID	ADF23294		
AC	ADF23294 standard; DNA; 5051 BP.		
XX	ADF23294;		
DT	12-FEB-2004 (first entry)		
DE	Recombinant MevB operon MBI.		
KW	isopentenyl pyrophosphate; mevalonate pathway;		
KW	isopentenyl pyrophosphate synthetase; isoprenoid; pharmaceutical;		
KW	nutraceutical; flavouring agent; agricultural pest control agent;		
KW	acetosuccinyl-CoA thiolase; acbB; Mevalonate bottom operon; MevB operon;		
KW	gene; ds1 isopentenyl pyrophosphate isomerase; idi;		
KW	dimethylallyl pyrophosphate; DMAPP; IPP; MBI.		
XX			
OS	Synthetic.		
OS	Escherichia coli.		
OS	Bacteria.		
XX			
PN	US2003148479-A1.		
XX			
PD	07-AUG-2003.		
XX			
PF	06-DEC-2001; 2001US-0006909.		
XX			
PR	06-DEC-2001; 2001US-0006909.		
PA	(KEAS/) KEASLING J.		
PA	(MART/) MARTIN V.		
PA	(PIRE/) PITERA D.		
PA	(KIMS/) KIM S.		
PA	(WITH/) WITHERS S T.		
PA	(YOSH/) YOSHIKUNI Y.		
PA	(NEWM/) NEWMAN J.		
PA	(KHE/) KHEBNIKOV A V.		
XX			
PI	Keasling J, Martin V, Pitera D, Kim S, Withers ST, Yoshikuni Y,		
PI	Newman J, Khebnikov AV;		
XX			
DR	WPI, 2004-020454/02.		
XX			
PT	Synthesizing isopentenyl pyrophosphate in a host microorganism, useful		
PT	for pharmaceutical purposes, comprises introducing into the microorganism		
PT	heterologous nucleic acid sequences coding for an enzyme in the		
PT	mevalonate pathway.		
XX			
PS	Example 1, SEQ ID NO 12; 40pp; English.		
XX			
CC	The invention describes a method of synthesizing isopentenyl		
CC	pyrophosphate (1) in a host microorganism. The method comprises		
CC	introducing into the host microorganism a plurality of heterologous		
CC	nucleic acid sequences each coding for a different enzyme in the		
CC	mevalonate pathway for producing isopentenyl pyrophosphate. The method is		
CC	useful in synthesizing isopentenyl pyrophosphate and the derived		
CC	isoprenoids. The isoprenoids may be used in pharmaceuticals,		
CC	nutraceuticals, flavouring agents or in agricultural pest control agents.		
CC	This sequence represents the MBI operon created from adding the B. coli		
CC	isopentenyl pyrophosphate isomerase gene idi to the MevB operon to ensure		
CC	sufficient production of dimethylallyl pyrophosphate (DMAPP) from		
CC	isopentenyl pyrophosphate (IPP), required to create the backbone		
CC	structure of all isoprenoids.		

XX	Sequence	5051 BP, 1486 A, 1064 C, 1122 G, 1379 T, 0 U, 0 Other;
Query Match	10.2%; Score 132; DB 12; Length 5051;	
Best Local Similarity	50.5%; Pred. No. 1,4e-24;	
Matches	473; Conservative 0; Mismatches 430; Indels 33; Gaps 5	
QY	3 GTCAAAAGCAATTGATGTCACCTGGAAAAAGCATTTCTTGCTGTGTGATTAATTTGGTTCTTGA	62
DB	1606 GTTGAGACCTTCAGTCCGCCAGGAAAGGCTTACTAGCTGGTGATATTATTTAGTTTAA	1661
QY	63 GCCAATTTATGATGCTTATGTCACAGCATTTGTCATACGAAATGCAAGCTTAACACC	122
DB	1666 TACAAAATATGAAAGCATTTGTATCGGATTAATCGGAAAGATGCAAGCTGTAGCCCATCC	1722
QY	123 AAAAGAAC-----CAGTTGAAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTTGC	176
DB	1126 TTAGCGTTCAATGCAAGGCTGTATAGTTTGAAGTGCCTGTGAAAAGTAAACAAATTTAA	1788
QY	177 AAACGAGATGGGAATATCAACATATCATCAAAATCAG--AGAAGCCAGAAAGTTCA	233
DB	1786 AGATGGGAGATGGCTGTACCAATTAAGTCTTAAAGTGGCTTCATCTCGTTTCGATVGG	1845
QY	234 GTACAGCATTAATCCATTTTATGAGGCAACTATATTCATCGTTTATGCTTATATTAACAC	293
DB	1846 CGAGTCTAAGAACCTTTTCATTTGAAAAAGTATCCCTAACGATTTACTCTACTTTAAACC	1905
QY	294 GACCAAGCATTTGATCTTGAAA-----TCATCATTTTACTCAGACCTGGATATCA	344
DB	1906 TAAATATGAGACGACTGTGCAATAGAACTTTGTTCTGTATTAATTTTCTGTATGATGC	1965
QY	345 TTCAACAAGAAATACTGAAACCAAGACATCTCGATATGAGAAAAAAACATTTCTTTTACA	404
DB	1966 CTACACTTCTCAGAGGATAGCGTTACCGAACATCGTGGCAACAGAAATTTGAATTTTCA	2025
QY	405 TTCTCGTGCATTTTACCGAAGTGGAAAAAACCAGATTAAGTTTCATCGCAGATTAAGTGC	464
DB	2026 TTGCGACAGAAATTTGAAGAAAGTTCCCAAAACAGGCTGGGCTCTCGCGCAGGTTTATGAC	2085
QY	465 AGTTGTGCACAAGTTTATTAATCCATTTTATCCC-----CAATGTATCAATAC	515
DB	2086 AGTTTAACTACAGCTTTGGCCTCCTTTTGTATCGAAGCTGGAAAAATATGTAGCA	2145
QY	516 GAATTAAGATTTTGTGCAACAAGTTGCAACAGATTTGCAATTTGATGCCAAAAAAGAT	575
DB	2146 AATTGAAAGATTTATCTAATTTTATGACACAAGTTGTCATATGTCACACTCAGGCTAAAT	2205
QY	576 AGAATCTGGGTTTGAATGTTGCACTGCAATTTTATGCTCATTTGATATATGAGATTTTCA	635
DB	2206 TGGAAAGCGGTTTATGATGAGCGGCGGACGATATGATCTATCAGATATAGAAATTTCC	2265
QY	636 GCCAGCTTTGTAATATGACGTTTCAAGGTTCTTAAAGATGATCTGAGAAAGTTCCAC	695
DB	2266 ACCGCAATTAATCTCTAATTTTGCAGATTT-----GGAAAGTCTACTTAACGCGATGA	2319
QY	696 AAGATTGAAAAAATGATTAAGAAAGTAACTGGAAATTCAAACATGAAGAATGTACATTAAC	755
DB	2320 ACTGAGCGCATTTGGTGTATGAAAGAAAGCTGGAATATTAAGATTAAGATTAACATTTAC	2379
QY	756 ATAGGGAATCAAGTTATTAATGGGTGACGTCAAGGGTGGCTGAGAAACCCAAATTTGT	815
DB	2380 TTGGGGAATTAATCTTTATGATGGGCAATTTAAGATGAGTTGCAAGAAACGTAATTAAC	2439
QY	816 ATCAAGATCTCAATGAGAAAAAGGAAAAAGCAAGAAAGAACTCTGTTGTATGACCA	875
DB	2440 CAGAAAGTAAAAAATTTGATATTCGATATGCGATATGCGAAAGCTTGAATTAATACGA	2499
QY	876 GCTTAATAGTCCAAATTTACAGTTTATGAGAAATTT 911	
DB	2500 ACTCGATCAATGCAAAATTTAGATTTATGATGAGACT 2535	

ID AD182322
 AC AD182322 standard; DNA; 5051 BP.
 XX AD182322;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Amorph-4,11-diene synthesis-related MBI operon DNA sequence Seg1012.
 XX
 KW amorph-4,11-diene synthesis; mevalonate pathway;
 KW isopentenyl pyrophosphate; amorph-4,11-diene synthase gene;
 KW pharmaceutical; nutraceutical; flavouring agent;
 KW agricultural pest control agent; epi-cedrol; MBI operon; ds.
 XX
 OS *Saccharomyces cerevisiae*.
 OS *Escherichia coli*.
 XX
 PN US2004005678-A1.
 XX
 PD 08-JAN-2004.
 XX
 PP 09-APR-2003; 2003US-00411066.
 XX
 PR 06-DEC-2001; 2001US-00006909.
 XX
 PA (KEAS// KEASLING J.
 PA (MART// MARTIN V.
 PA (PITTE// PITTERA D.
 PA (WITTH// WITHERS S T.
 PA (NEWM// NEWMAN J.
 XX
 PI Keasling J, Martin V, Pittera D, WITHERS ST, Newman J;
 XX
 WP1; 2004-120864/12.
 XX
 XX Synthesizing amorph-4,11-diene in a host cell, useful as
 PT pharmaceuticals, comprises introducing nucleic acid sequences, each
 PT coding for a different enzyme in the mevalonate pathway for producing
 PT isopentenyl pyrophosphate.
 XX
 PS Example 1, SEQ ID NO 12, 75bp; English.
 XX
 CC This invention relates to a novel method of synthesizing amorph-4,11-
 CC diene in a host microorganism which comprises introducing into the host
 CC microorganism heterologous nucleic acid sequences, each coding for a
 CC different enzyme in the mevalonate pathway for producing isopentenyl
 CC pyrophosphate, and introducing into the host microorganism a DNA fragment
 CC coding for an optimised synthetic amorph-4,11-diene synthase gene. The
 CC nucleic acid sequences, DNA fragments, vectors and host cells are useful
 CC for synthesizing isopentenyl pyrophosphate, amorph-4,11-diene and epi-
 CC cedrol used as pharmaceuticals, nutraceuticals, flavouring agents, or
 CC agricultural pest control agents. The present sequence is that of an
 CC operon which may be used in the method of the invention.
 XX
 XX Sequence 5051 BP; 1486 A; 1064 C; 1122 G; 1379 T; 0 U; 0 Other;

Query Match	10.2%	Score 132	DB 12	Length 5051
Best Local Similarity	50.5%	Pred. No. 1.4e-24		
Matches 473	Conservative	0	Mismatches 430	Indels 33
			Gaps	5
Qy	3	GTCAAAAGCATTTTGTGTCACCTGTGAAAAGCATTTCTCTGTGTGATATTTTGGTCTTGA	62	
Db	1606	GTTGAGAGCCTTCAGTGCCTCCCGGAAAAGCGTTACTAGCTGTGATATTAGTTTGA	1665	
Qy	63	GCCATTATATGATGCTTATATGTGACAGCATTTGTCATCAGATGCAATGCAGTTATAACCC	122	
Db	1666	TACAAAATATGAAAGCATTTGTATGTGGATTTAACGGCAAGAAAGCATGCTGTAAGCCATCC	1725	
Qy	123	AAAAGGAAAC-----CAGTTGAAAGATCTAGATCAAAATTTCTGACCCCAATTTGCG	176	
Db	1726	TTACGCTTCATTGCGAAGGCTCTGATTAAGTTGAAGTGCCTGTGAAAAGTAACAATTTTA	1785	
Qy	177	AAACGGAGATGGGATATCACTATTCATCAAAATACAG--AGAAAGCCAGAGAGTTCA	233	

Db	1786	AGATGGGGAGTGGCTGTACCATATTAAGCTCTAAAGTGGCTTCAATCTCTGTTTGATAGG	1845
Oy	234	GTCAACGATTAATCCATTTTATGAGGCACTATAATCAATCGTTTATGCTTAATTAAC	293
Db	1846	CGGATCTTAAGAACCTTTCATTTAGTAAAAAGTTATGCTTAACGTAATTAATGCTACTTTAAAC	1905
Oy	294	GACCGAAGCATTTGATCTTGAAA-----TCATCATTTTACTCAGACCCCTGGATATCA	344
Db	1906	TAACTAGGACGACTACTGCAATTAAGAACTTGTTGTTATTAATATTTTCTTGATGATGC	1965
Oy	345	TTCAACAGAAAGTACTGAAACCAAGACATCTCGAATGAGAAAAACATTTCTTTTACA	404
Db	1966	CTACACATCTCAGGAGGATAGCTTTACGAAATCATGATGGCAACAGAAAGATTGAAGTTTCA	2025
Oy	405	TTCTTCGTCCATTACCGAATGAAAAAGACCGGATTAGTTTATCGGCAGAGTTATGTC	464
Db	2026	TTCCACAGAAATTAAGAAAGTTCCCAAAACAGGGCTGGCTCCTCGGAGGTTTAAGTAC	2085
Oy	465	AGTTGTTGCCCAAGTTTATTAATCCCATTTTATCCC-----CAATGTTATCAGTAC	515
Db	2086	AGTTTAACTACACACTTTGGCTCTCTTTTGTATCGAAGCTGGAAATATATGAGACA	2145
Oy	516	GAATTAAGATATTTTGCACAAGTTGACAGATTGCAATTTGATGCCCCAAAAAAGAT	575
Db	2146	ATATAGAGAAAGTTATTCATTAATTTAGACAAAGTTGCTCATTTGCAAGCTCAGGGTTAAAT	2205
Oy	576	AGGATCTGGGTTTATGTTGCAATCGCAATTTATGATCTGATTTGATATGAAAGTTTCA	635
Db	2206	TGGAAGCGGGTTTATGATGAGCGGCGGACGATATGAGATCTATACAGATATGAAAGTTCCC	2265
Oy	636	GCCAGCTTTGATTAATGACGTGTTCAGGTTCTAGAAAGTGATCCTGAGAACTTCCCCAC	695
Db	2266	ACCGGCAATTAATCTCTAATTTGGCAAGATTT-----GGAAGTGTACTTATCGGCAATGA	2319
Oy	696	AGAGTTGAAAAAATTTGATTTGAAAGTAATCTGGGAATTCAAACATGAAGATGTACATTAC	755
Db	2320	ACTGCGCGATTTGGTGTGATGAAAGACCTGGAAATTTATACGATTAAGTAACCATTTAAC	2379
Oy	756	ATACGGATCAAGTTATTAATGSGTGACGTCAAGGGTGGCTCAGAAAACCCCAATTTGAT	815
Db	2380	TTCCGGATTAACCTTATGATAGGCGCATTTAAGATATGTTCCAGAAACAGTAAACATGCT	2439
Oy	816	ATCAACGACTCTCAATGAGAAAAAGAAAAGCCAGAAAGAAAGCTCTGTTGTATGACCA	875
Db	2440	CCAGAAGGTAAAAAATGTGTATGATTCGCATATCCCAAGAAAGCTTGAATAATATATACGA	2499
Oy	876	GCTTAATATGTCCAATTTACAGTTTATGAAGGAATTT 911	
Db	2500	ACTGATCATGCAAAATTCATGATTTTATGATGAGTGAAGCT 2535	

```

RESULT 17
ADef23295
ID ADef23295 standard, DNA, 5963 BP.
XX
XX ADef23295;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
XX Recombinant MevB operon MBIS.
DE
XX
XX isopentenyl pyrophosphate: mevalonate pathway;
KM isopentenyl pyrophosphate synthetase; isoprenoid; pharmaceutical;
KM nutritional; flavouring agent; agricultural pest control agent;
KW acetoacetyl-CoA thiolase; acOB; mevalonate bottom operon; MevB operon
gene; ds; farnesyl pyrophosphate synthase; lspa; MBIS.
XX
XX Synthetic.
OS Escherichia coli.
OS Bacteria.
XX
XX US2003148479-A1..
PN

```

```

XX 07-AUG-2003.
PD
XX 06-DEC-2001; 2001US-00006909.
PF
XX 06-DEC-2001; 2001US-00006909.
PR
XX (KEAS/) KEASLING J.
PA (MART/) MARTIN V.
PA (PITERA/) PITERA D.
PA (KIMS/) KIM S.
PA (WITHERS S T/) WITHERS S T.
PA (YOSH/) YOSHIKUNI Y.
PA (NEWM/) NEWMAN J.
PA (KHLEB/) KHLERBNIKOV A V.
PI Keasling J, Martin V, Pitera D, Kim S, Withers ST, Yoshikuni Y;
PI Newman J, Khlernikov AV;
XX
XX WPI; 2004-020454/02.
DR
XX
XX
XX PT Synthesizing isopentenyl pyrophosphate in a host microorganism, useful
XX PT for pharmaceutical purposes, comprises introducing into the microorganism
XX PT heterologous nucleic acid sequences coding for an enzyme in the
XX PT mevalonate pathway.
XX
XX PS Example 1; SEQ ID NO 13; 40bp; English.
XX
XX CC The invention describes a method of synthesizing isopentenyl
XX CC pyrophosphate (I) in a host microorganism. The method comprises
XX CC introducing into the host microorganism a plurality of heterologous
XX CC nucleic acid sequences each coding for a different enzyme in the
XX CC mevalonate pathway for producing isopentenyl pyrophosphate. The method is
XX CC useful in synthesizing isopentenyl pyrophosphate and the derived
XX CC isoprenoids. The isoprenoids may be used in pharmaceuticals,
XX CC nutraceuticals, flavouring agents or in agricultural pest control agents.
XX CC This sequence represents the MBIS operon comprising the E. coli farnesyl
XX CC pyrophosphate synthase gene isppa added to the MB1 operon which contains
XX CC the Merv operon plus the isopentenyl pyrophosphate isomerase gene to
XX CC direct the product of the mevalonate pathway operon.
XX
XX SQ Sequence 5963 BP; 1700 A; 1305 C; 1374 G; 1584 T; 0 U; 0 Other;

Query Match      10 2%; Score 132; DB 12; Length 5963;
Best Local Similarity 50.5%; Pred. No. 1.5e-24;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTGAGAAAGCATTTCTGCTGGTGGATATTTGGTCTTGA 62
DB 1606 GTTGAGAGCCTTCAAGTCCCGCCAGGAAAGCCTTACTAGCTGGATATTTAGTTTGA 1665
QY 63 GCCAATTTATGATGCTTATGTCAGCATTTGTCATCAGCATGTCAGTATTAACAC 122
DB 1666 TACAAAAATATGAAGCATTTGTATGCGATTAATCGCAAGAAATGCATGCTGATGCCATCC 1725
QY 123 AAAAGAAC-----CAGTTGAAGAAATCTAGAAATCAAAATTTTTCACCCCAATTGGC 176
DB 1726 TTACGCTTCATTGCAAGGCTCTGATTAAGTTGAATGCGGTGAGAAAGTAAACAATTAA 1785
QY 177 AAACGAGAAATGGAATATCAATATCATATCAATCAAG---AGAAGCCAGAGAAAGTTCA 233
DB 1786 AGATGGGGAGTGGCTGTACATATTAAGCTCTAAAGTGGCTTCACTCTGTTTGCATAGG 1845
QY 234 GTACAGGCATTAATTCATTTTATGAGCAACTATATTCATCGTTTATAGCTTATATTAACA 293
DB 1846 CGGATCTAAGAACCTTTTCATTGAAAGAAAGTATTCCTAAGCATTTAGCTATTAACCC 1905
QY 294 GACCGAAGCATTTGATCTTGAAG-----TCAATCTTATACAGACCTCGATATCA 344
DB 1906 TAACTGACGACATCTGCAATAGAAACTGTTTGGTATATTAATTTTCTCTGATGATGC 1965
QY 345 TTCAACAGAAATATCTGAACCAAGACATCTCGAATGAGAAAGAAACATTTCTTTACCA 404

```

```

DB 1966 CTACCATTTCTGAGAGATAGCGTTACCGAAACATCGTG3CAACAGAGATGAGTTTCA 2025
QY 405 TTCTGTCGATTCATTACCCAGATGAGAAAGACCGGATTTAGTTCATGCGACAGATTAGTGC 464
DB 2026 TTGCGACAGAAATGAGAGATGTCCTCCAAAACAGGGCTGGGCTCTCGGACAGGTTTAGTAC 2085
QY 465 AGTTGTTCACCAAGATTATTTATCCCATTTTATCC-----CAATGTTATCAGTAC 515
DB 2086 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATGCACTCGAAGAAATTAATGATGACAA 2145
QY 516 GAATTAAGATATTTTGCACAAAGTTGCACAGATTGCACTGTTATGCCCCAAAAAGAT 575
DB 2146 ATATAGAGAAAGTTATTCATAATTTAGCACAAGTTGCTCATTTGCAAGCTCAGGTTAAAT 2205
QY 576 AGGATCTGGGTTTGAATGTTGCACTGCAATTTATGCTGTGATTAATGATTAAGATTCA 635
DB 2206 TGGAAAGCGGTTTGTATGAGCGCGGACATTAATGATCTATACATATTAAGAAATTTCC 2265
QY 636 GCCAGCTTTGATTAATGACGTGTTTCAAGTTCTAGAAAGTATCTGGAAGATTCCCCAC 695
DB 2266 ACCGCAATTATCTCTAATTTGCCAGATTT-----GGAAGTCTACTTACGGCAGTAA 2319
QY 696 AGAGTTGAAAAAATTTGATTGAAAGTAACTGCGAATTCAAACATGAAAGATGATACATACC 755
DB 2320 ACTGCCGATTTGGTGTATGATGAGAAAGACTGGAATATTTACATTAAGTAACCATTTAAC 2379
QY 756 ATACGAATCAAGTTATTAATGAGTGAAGTCAAGAGGTGGCTCAGAAACACCAATTTGTT 815
DB 2380 TTGCGGATTAATCTTATGATGAGCGCATTTAAGATATGTTCAAGAAACAGTAAACTGTT 2439
QY 816 ATCAAGAGTACTCCATGAGAAAGAAAGAAAGCCAGAAAGAAAGCTTGTGTATGACCA 875
DB 2440 CCAGAGAGTAAAGAAATTTGATGATGATGCGATATGCGAATGCGAAGAAAGTTGAAATATATACGA 2499
QY 876 GCTTAATATGAGCCAAATTTACAGTTTATGAAGAAAT 911
DB 2500 ACTGATCATGCAAAATTTCTAGATTTATGATGACT 2535

RESULT 18
ADI82323
ID ADI82323 standard; DNA; 5963 BP.
AC ADI82323;
XX
XX 22-APR-2004 (first entry)
DT
XX
XX Amorpha-4,11-diene synthase is-related MBIS operon DNA sequence SegID13.
XX
XX amorpha-4,11-diene synthase; mevalonate pathway;
XX isopentenyl pyrophosphate; amorpha-4,11-diene synthase gene;
XX pharmaceutical; nutraceutical; flavouring agent;
XX agricultural pest control agent; epi-cedrol; MBIS operon; ds.
XX
XX Saccharomyces cerevisiae.
OS
XX Escherichia coli.
XX
XX US2004005678-A1.
XX
XX 08-JAN-2004.
XX
XX 09-APR-2003; 2003US-00411066.
XX
XX 06-DEC-2001; 2001US-00006909.
XX
XX (KEAS/) KEASLING J.
XX (MART/) MARTIN V.
XX (PITERA/) PITERA D.
XX (WITHERS S T/) WITHERS S T.
XX (NEWM/) NEWMAN J.
XX
XX Keasling J, Martin V, Pitera D, Withers ST, Newman J;

```

DR WPI; 2004-120864/12.
 XX Synthesizing amorpho-4,11-diene in a host cell, useful as
 PT pharmaceuticals, comprises introducing nucleic acid sequences, each
 PT coding for a different enzyme in the mevalonate pathway for producing
 PT isopentenyl pyrophosphate.
 XX
 PS Example 1; SEQ ID NO 13; 75bp; English.
 XX
 CC This invention relates to a novel method of synthesizing amorpho-4,11-
 CC diene in a host microorganism which comprises introducing into the host
 CC microorganism heterologous nucleic acid sequences, each coding for a
 CC different enzyme in the mevalonate pathway for producing isopentenyl
 CC pyrophosphate, and introducing into the host microorganism a DNA fragment
 CC coding for an optimised synthetic amorpho-4,11-diene synthase gene. The
 CC nucleic acid sequences, DNA fragments, vectors and host cells are useful
 CC for synthesizing isopentenyl pyrophosphate, amorpho-4,11-diene and epi-
 CC cedrol used as pharmaceuticals, nutraceuticals, flavouring agents, or
 CC agricultural pest control agents. The present sequence is that of an
 CC operon which may be used in the method of the invention.
 XX
 SQ Sequence 5963 BP; 1700 A; 1305 C; 1374 G; 1584 T; 0 U; 0 Other;

Query Match 10.2%; Score 132; DB 12; Length 5963;
 Best Local Similarity 50.5%; Pred. No. 1.5e-24;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAGACATTTCTGTGCGATATTTGGTTCTTGA 62
 DB 1606 GTTGAGCGCTTCAGTGGCCCGAGGAGAACGTTACTAGCTGTGATTTAGTTTAA 1665
 QY 63 GCCAATTTATGATGCTTATGTGACAGCATTTGCATCAGCATGATGATTAACACC 122
 DB 1666 TACAAATATGATGATTTGTAGTGGATTAATCGCAAGAAATGCATGCTGAGCCATCC 1725
 QY 123 AAAAGAAC-----CAGTTGAAAGAAATCTAGAAATCAAAATTTCTTCAACCCCAATTGTC 176
 DB 1726 TTACGCTTCATTTGCAAGGCTCTGATAGTTGAAGTCGTGAAAGATTAACAAATTGA 1785
 QY 177 AAACGAGAAATGGGAATATACATATCATCAATTAACAG--AGAACCCAGAGAAAGTTCA 233
 DB 1786 AGATGGGAGTGGCTGTACATATTAATGCTTAAAGTGGCTTCATTTCCGTTTCGATAG 1845
 QY 234 GTACAGCATTAATTCATTTTATGAGGCACTATATTCATCGTTTATGTTATATTAAC 293
 DB 1846 CGGATCTAAGAACCTTTTCATTTGAAAGATTAATCGTAAAGATTTAGTACTTTAAAC 1905
 QY 294 GACCGAAGCATTTGATCTTGAAG-----TCATCATTTTATCTGAGACCTGATATCA 344
 DB 1906 TAAACATGACGACTACTGCAATAGAAATCTGTTGTTATGATTTCTCTGATGATGC 1965
 QY 345 TTCAACAGAAAGTACTGAAACCAAGACATCTCGAATGAGAAAGAAATCTTTCTTACCA 404
 DB 1966 CTACCATTTCTCGAGAGATAGGTTAACCCAGACATCGTGGCAACAGAAAGTTAGTTTCA 2025
 QY 405 TTCTCGTCGATTTACCGAAGTGGAAAGACCGGATATTAGTTTCGCGAGATTAAGTGC 464
 DB 2026 TTTCGACAGAAATTTGAAGAAATTCCTCAAAACAGGGCTGCTCGGACGGTTTACTGC 2085
 QY 465 AGTTGTTGCGACAGATTTATTTATCCCATTTTATCC-----CAATGTTATCAGTAC 515
 DB 2086 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATCGACCTGGAAGAAATATGTAGACA 2145
 QY 516 GAATTAAGATATTTTGCACACAGTTCGACAGATTCGACATTTGTATGCCCAAAAAAGAT 575
 DB 2146 ATATGAGAGAGATTTATTCATTAATTTAGCAAAAGTTGCTCTTTCATTCAGCTCAGGTA 2205
 QY 576 AGGATCTGGGTTTGAATGTTGCAAGTCAATTTATGCTGATGATATAGAAAGATTTCA 635
 DB 2206 TGGAAAGCGGTTTGAATGTTGCGCGGCGGACATATGATCTATCAATATAGAAAGATTTCC 2265
 QY 636 GCCACCTTTGATTAATGACGTGTTTCAGGTTCTAGAAAGTGAATCTGGAAGATTTCCCCAC 695

DB 2266 ACCGCGATTAATCTCTTAATTTGCCAGATATT-----GGAACTGCTACTTACGGCATTA 2319
 QY 696 AGAGTTGAAAAAATTTGATTAAGAAATGACTGGAAATTTCAACATGAAGATGTATAC 755
 DB 2320 ACTGGCGCATTTTGTGTTATGAAGAACTGGAATTTTCAGATTAAGTAACATTTTACC 2379
 QY 756 ATACGGAATCAAGTTATTAATGTTGTTGACGTCAAGGTTGCTCAGAAACCCCAATTGGT 815
 DB 2380 TTCCGGATTTAACTTTATGATGATGGCGCATATTAAGATGGTTTCAGAAACAGTAAACTGGT 2439
 QY 816 ATCAGAGTACTCCAAATGAAAAAGCAAGAAAGCTCTGTTGTATGACCA 875
 DB 2440 CCAGAAAGTAAAAAATTTGATGATTTCCGATATGCCAGAAAGCTTGAAAAATATTAACAGA 2499
 QY 876 GCTTATAGTGGCAATTTTACAGTTTATGAAGGATTT 911
 DB 2500 ACTGCATCATGCAAAATTTCAAGTTTATGATGATGACT 2535

RESULT 19

AAD31026
 ID AAD31026 standard; DNA; 7681 BP.

XX AAD31026;
 XX 29-AUG-2003 (revised)
 DT 31-MAY-2002 (first entry)
 XX
 DE Operon D DNA encoding mevalonate pathway.

XX Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
 KW isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
 KW transgenic plant; yeast; phosphomevalonate kinase; HMGR; HMGs; AACT;
 KW mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MMD;
 KW acetoacetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
 KW MVK; PMK; HMG-CoA reductase; ds.

OS Saccharomyces cerevisiae.
 OS Arabidopsis thaliana.
 OS Streptomyces sp.
 OS Chimeric.

PN WO200210398-A2.
 XX
 PD 07-FEB-2002.

PF 31-JUL-2001; 2001WO-US024037.

PR 31-JUL-2000; 2000US-0221703P.

PA (HANN/) HAHN P M.
 PA (KUEH/) KUEHNLS A R.

PI Hahn FM, Kuehnle AR;

DR WPI; 2002-217122/27.

XX Use of specific genes of mevalonate and isoprenoid biosynthetic pathways,
 PT for providing a cell with herbicide or antibiotic resistance, and for
 PT providing transformed cells having increased isoprenoid production.

PS Claim 77; Page 141-145; 193bp; English.

XX The invention relates to the use of specific genes of the mevalonate and
 CC isoprenoid biosynthetic pathways and inactive gene sites (pseudogene).
 CC Genes of the invention are used to enhance biosyntheses of isopentenyl
 CC diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid
 CC pathway derived products in the plastids of transgenic plants and
 CC microalgae, for producing herbicide or antibiotic resistant transgenic
 CC plants and microalgae, for providing transformed cells with increased
 CC isoprenoid production compared to non-transformed cells, and for
 CC providing a cell with an inserted polynucleotide sequence encoding one or
 CC more products of interest. The present sequence is operon D DNA encoding

CC the entire mevalonate pathway. This operon contains *S. cerevisiae* orfs
CC encoding phosphomevalonate kinase (PMK), mevalonate kinase (MVK),
CC mevalonate diphosphate decarboxylase (MDD), acetoacetyl thiolase (AAT)
CC and *A. thaliana* orfs encoding 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-
CC-CoA) synthase (HMS) and Streptomyces sp. Cl197 of encoding HMG-CoA
CC reductase (HMGR). (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 7681 BP; 2127 A; 1691 C; 1837 G; 2026 T; 0 U; 0 Other;

Query Match	10.2%	Score 132,	DB 6;	Length 7681;
Best Local Similarity	50.5%;	Pred. No. 1.6e-24;		
Matches 473; Conservative	0;	Mismatches 430;	Indels 33;	Gaps 5

QY	3	GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGCTGTGGATTTTGGTTCTTGA	62
Db	43	GTGAGAGCCCTTAGTGCCCCAGGAAAAGGTTACTAGCTGGTGGATTTTATGTTTGA	102
QY	63	GCCAAATTATGATGCTTATGTGACAGCATTTGATCATCAGATGATGCAGTTTAAACACC	122
Db	103	TACAAATATGAAGCATTTTGTATGTCCGATTTATCGCAGAATGCATCTGAGCCATCC	162
QY	123	AAAAGAAC-----CAGTTGAAGAATCTAGATCAAAATTTCTTACCCCAATTTGC	176
Db	163	TTACGGTTCAATGCCAAGGGCTGTATTAAGTTTGAAGTCCGTGTGAAAAGTAAACATTTAA	222
QY	177	AAACGAGATGSGGAATATACATATCATCAATTAACG---AGAGCCCAAGAAAGTTCA	233
Db	223	AGATGGGAGTGGCTGTACATTAATTAAGTCTTAAAGTGCTTCAATCTCTTTGCAATAGG	282
QY	234	GTTCACGCATTAATTCATTTTAAAGGCACATATATCATGTTTAACTTATATTCACCC	293
Db	283	CGGATCTAAGAACCCCTTTCATTTAATAAAGTTATCGCTAAGTATTAAGTACTTAAACC	342
QY	294	GACCGAAGCATTTGATCTTGAAG-----TCATCATTTACTCAGACCCCTGATATCA	344
Db	343	TAACATGAGACGACTCTGCATATGAAACTTGTTGTTATTAATTAATTTCTCTATGATGC	402
QY	345	TTTCACAAAGATATCTGAAACCAAGCATCTCGAATGAGAAAACATTTCTTTAACCA	404
Db	403	CTACCATCTTCAGAGAGATAGCGTTACCAACATCGTGGCACAAGAAATTAAGTTTCA	462
QY	405	TTCTCGTGCATTTACCGAAGTGGAAAAGCCGATTAGTTTATCGGCAGATTAAGTGC	464
Db	463	TTTCGACAGAAATTGAAGAAGTTCCTCAAAACAGGGCTGGGCTCTCGGCAAGTTTAATGCAC	522
QY	465	AGTTGTGACACAAGTTTATTTATCCCATTTTATCCC-----CAATGTTATCAGTAC	515
Db	523	AGTTTAACTACAGCTTTGGCCCTCTTTTGTATGGAAGCTGGAAAATATATGACACA	582
QY	516	GAATTAAGATTTTTCGACACAAGTTGCAAGATTGCAATGTTTATGCCCAAAAAAAGAT	575
Db	583	ATATAGAGAAAGTTATTCATAATTTTACACAAGTGTCTCATTTGCAAGCTCAGGGTAAAT	642
QY	576	AGATCTGGGTTTGATAGTTGCAACTGCATTTATAGTGTGATTTATAGAAGATTTCA	635
Db	643	TGGAAGCGGGTTTGAATGTAACGGCGGACACATATGATCTATCAGATTAAGAAGATTTCC	702
QY	636	GCCACCTTGTAAATATGACGTGTTTCAAGTTCTAGAAGATGATCTGAGAAGTTCCCAAC	695
Db	703	ACCCGCATTAAATCTCTAAATTTCCAGATATTT-----GGAAGTGCTACTTACGGCAGTAA	756
QY	696	AGAGTTGAAAAAATGATTTGAAGATPACCTGGGAATTCACAACATGAAAGATGTACATTTAC	755
Db	757	ACTGGCGCATTTGTTGATGAGAAAGACCTGGAATATTAAGATTAAGAACCATTAATTAAC	816
QY	756	ATACGGAATCAAGTTATTAATGGGTGACGTCGAAGGTGGCTCAGAAACACCCAAATTTGT	815
Db	817	TTTCGGATTTAACTTTATGATGGGCGAATTTAAGAATGTTTCAGAAACGTAATTAACGTGT	876
QY	816	ATCAGAGATCACTCAATGGAAAAAAGCAAGAAAGCAAGAAACCTCTGTTGTATGACCA	875
Db	877	CCAGAGGTTAAAAATTTGGTATGATTTCCGATATGCCAGAAACCTTGAATAATTTATACGA	936

Qy 876 GCTTAATAGTCCCAATTACAGTTTATGAGGAATT 911
| | | | | | | | | | | | | | | | | |
Db 937 ACTGCATATGCAATTCTAGATTATGATGAGACT 972

RESULT 20
AAD31023
ID AAD31023 standard; DNA; 7693 BP.

DT	29-AUG-2003	(revised)
DT	31-MAY-2002	(first entry)

Operon A DNA encoding mevalonate pathway.

KM Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP
KM isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
transgenic plant; yeast; phosphomevalonate kinase; HMGR; HMGS; ACT;
KM mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MDD
KM acetoacetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
KM MK; PK; HMG-CoA reductase; ds

OS *Saccharomyces cerevisiae*.
OS *Arabidopsis thaliana*.
OS Chimeric.

PN WO200210398-A2.

PD 07-FEB-2002.

PF 31-JUL-2001; 2001WO-US024037.

PR 31-JUL-2000; 2000US-0221703P.

PA (HAHN/) HAHN F M.

PI Hahn FM, Kuehnle AR;

WPI; 2002-217122/27.

PT Use of specific genes of mevalonate and isoprenoid biosynthetic pathways, for providing a cell with herbicide or antibiotic resistance, and for providing transformed cells having increased isoprenoid production.

PS Claim 77; Page 127-131; 193pp; English.

The invention relates to the use of specific genes of the mevalonate and isoprenoid biosynthetic pathways and inactive gene sites (pseudogenes). Genes of the invention are used to enhance biosynthesis of isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid pathway derived products in the plastids of transgenic plants and microalgae, for producing herbicide or antibiotic resistant transgenic plants and microalgae, for providing transformed cells with increased isoprenoid production compared to non-transformed cells, and for providing a cell with an inserted polynucleotide sequence encoding one or more products of interest. The present sequence is operon A DNA encoding the entire mevalonate pathway. This operon contains *S. cerevisiae* orf encoding phosphomevalonate kinase (PMK), mevalonate kinase (MWK), mevalonate diphosphate decarboxylase (MDD), acetatecoyl thiolase (AACO) and *A. thaliana* orf encoding 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) synthase (HMGS) and HMG-CoA reductase (HMGR). (Updated on 29-AUG-2003 to standardise OS field)

SQ Sequence 7693 BP; 2212 A; 1526 C; 1794 G; 2161 T; 0 U; 0 Other;

Query Match	10.2%;	Score 132;	DB 6;	Length 7693;
Best Local Similarity	50.5%;	Pred. No. 1.6e-24;		
Matches 473; Conservative	0;	Mismatches 430;	Indels 33;	Gaps 5;

3 GTCAAAAGCATTTAGTGCACCTGAAAAGCATTTCTTGCTGATATTTGTTCTGA 62

Db 3786 GTTAGAGCTTCAGTCCCGGAGGAAACGTTACTAGCTGGTGAATTTAGTTTGA 3855
 QY 63 GCCAATTTATGATGCTTATGATGACAGCATTTGTCATCAGCAATGACATGATATAACACC 122
 Db 3856 TACAAATATGAGCATTTGTAGTGGATATCCGCAAGAAATGCAATGCTGTAGCCCATCC 3915
 QY 123 AAAAGAAC-----CAGTTGAAAGAAATCTAGAAATCAAAATTTCTTCAACCCCAATTTGC 176
 Db 3916 TTACGGTTCATTTGCAAGGGTCTGATTAAGTTGAAGGCGGTGAAAGTAAACAAATTTAA 3975
 QY 177 AAACGAGAAATGGGAAATATACATATCATCAATACAG--AGAACCCAGAGAAATTGA 233
 Db 3976 AGATGGGAGTGGCTGTACCATATTAAGTCTTAAAGTGGCTTCATTCCTGTTTCATAGG 4035
 QY 234 GTACGCAATTAATTCATTTTATGAGCAACTATTTATTCATGCTTTTATGATTTATTCACCC 293
 Db 4036 CGGATCTAAGAACCTTTTCATTTGAAAAGTTATCGTACGATATTTAGTACTTTAAACC 4095
 QY 294 GACGAGCATTTGATCTTTGAAA-----TCATCATTTTACTGAGACCTGGAATATGA 344
 Db 4096 TTAACATGACGACTACTGCAATGAAAATTGTTCCGTTATGATTAATTTCTGATGATGTC 4155
 QY 345 TTCACAAAGATGACTGAAAACCAAGACATCTCTGCAATGAGAAAACCTTTCTTTACCA 404
 Db 4156 CTACCATTTCTGAGAGATAGCGTTAACGACATCGTGGCAACAGAAATTTAGATTTTCA 4215
 QY 405 TTCTGTCGTCATTTACCGAAGTGAAGAAACCGGATATAGTTATCGGACGAGATTAAGTGTTC 464
 Db 4216 TTTCGACAGAAATTTGAAGATTTCCCAAAACAGGGCTGCGCTCTCGGAGTTTATGTCAC 4275
 QY 465 AGTTGTGGCAGCAAGTTTATTTATCCCATTTTATCC-----CAATGTTATCAGTAC 515
 Db 4276 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATGAGCATGGAATAATATGATGACAA 4335
 QY 516 GAATTAAGATATTTTGCACACCTTGCACAGATTTGCACTGTTATGCCCCAAAAGAT 575
 Db 4336 ATATGAGAGAGTTATTCATTAATTTGACACAGTTGCTCATTTGTCAAGCTCAGGGTAAAT 4395
 QY 576 AGATCTGGGTTTGAATGTTGCAACGTCATTTATGCTGATTTGATATAGAAATTTCA 635
 Db 4336 TGGAGACGGGTTGATGTAAGCGGCGGACATATGATTTATCAGATATAGAAATTTCC 4455
 QY 636 GCCACCTTTGATTAATGACGTTTCAAGTTCTAGAAAGTATCTGAGAAATTTCCCCAC 695
 Db 4456 ACCCGCATTAATCTCTAATTTGCCAGATAT-----GGAAGTGTACTTACGGCAGTAA 4509
 QY 636 AGATTTGAAAAAATTTGATGAAAATGTAATCTGGAATTTCAAACTGAAAGTGTACATTACC 755
 Db 4510 ACTGCGCATTTGTTGATGAAAGAAAGACTGGAATATTAAGATTTAAAGTAAACCATTTTACC 4569
 QY 756 ATACGAATTCAGATTATTAATGGGTGAGCTCAAGGGTGGCTGAGAAACCCAAATTTGGT 815
 Db 4570 TTCCGGAATTTACTTTATGATGAGGCGATATTAAGAAATGTTCAAGAAACGTAATCTGGT 4629
 QY 816 ATCAAGACTACTCCATGAGAAAAGAAAAGCCAGAGAAAGCTGTGTTGTATGACCA 875
 Db 4630 CCAGAAAGTAAATTTGTTGATGATTTCCGATATGCGAATGCGAAGAACTGTAATATATACGA 4689
 QY 876 GCTTAATAGTCCCAATTTACAGTTATGAGAAATTT 911
 Db 4690 ACTCGATCATGCAAAATTTAGATTTATGATGATGACT 4725

RESULT 21

ADD31024 standard; DNA; 7695 BP.

ADD31024;

AC AAD31024;
 XX 29-AUG-2003 (revised)
 DT 31-MAY-2002 (first entry)
 XX
 DE Operon B DNA encoding mevalonate pathway.

XX Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
 KW isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
 KW transgenic plant; yeast; phosphomevalonate kinase; HMG-R; HMG; AAT;
 KW mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MVD;
 KW acetoacetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
 KW MVK; PMK; HMG-CoA reductase; ds.
 XX Saccharomyces cerevisiae.
 OS Arabidopsis thaliana.
 OS Chimeric.
 PN WO200210398-A2.
 XX
 PD 07-FEB-2002.
 XX
 PP 31-JUL-2001; 2001WO-US024037.
 XX
 PR 31-JUL-2000; 2000US-0221703P.
 XX
 PA (HAHN/) HAHN F M.
 PA (KUEH/) KUEHNLE A R.
 XX
 PI Hahn FM, Kuehnle AR.
 XX
 DR WPI; 2002-217122/27.
 XX
 PT Use of specific genes of mevalonate and isoprenoid biosynthetic pathways,
 PT for providing a cell with herbicide or antibiotic resistance, and for
 PT providing transformed cells having increased isoprenoid production.
 PS
 PS
 PS Claim 77; Page 132-136; 193pp; English.
 CC The invention relates to the use of specific genes of the mevalonate and
 CC isoprenoid biosynthetic pathways and inactive gene sites (pseudogene).
 CC Genes of the invention are used to enhance biosynthesis of isopentenyl
 CC diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid
 CC pathway derived products in the plaids of transgenic plants and
 CC microalgae, for producing herbicide or antibiotic resistant transgenic
 CC plants and microalgae, for providing transformed cells with increased
 CC isoprenoid production compared to non-transformed cells, and for
 CC providing a cell with an inserted polynucleotide sequence encoding one or
 CC more products of interest. The present sequence is operon B DNA encoding
 CC the entire mevalonate pathway. This operon contains S. cerevisiae orfs
 CC encoding phosphomevalonate kinase (PMK), mevalonate kinase (MVK),
 CC mevalonate diphosphate decarboxylase (MVD), acetoacetyl thiolase (AAT)
 CC and A. thaliana orfs encoding 3-hydroxy-3-methylglutaryl-coenzyme A (HMG
 CC -CoA) synthase (HMG-S) and HMG-CoA reductase (HMG-R). (Updated on 29-AUG-
 CC 2003 to standardise OS field)
 SQ Sequence 7695 BP; 2214 A; 1524 C; 1793 G; 2164 T; 0 U; 0 Other;
 Query Match 10.2%; Score 132; DB 6; Length 7695;
 Best Local Similarity 50.5%; Pred. No. 1.6e-24;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
 QY 3 GCCAATTTATGATGCTTATGATGACAGCATTTGTCATCAGCAATGACATGATATAACACC 62
 Db 28 GTTAGAGCTTCAGTCCCGGAGGAAACGTTACTAGCTGGTGAATTTAGTTTGA 87
 QY 63 GCCAATTTATGATGCTTATGATGACAGCATTTGTCATCAGCAATGACATGATATAACACC 122
 Db 86 TACAAATATGAGCATTTGTAGTGGATATCCGCAAGAAATGCAATGCTGTAGCCCATCC 147
 QY 123 AAAAGAAC-----CAGTTGAAAGAAATCTAGAAATCAAAATTTCTTCAACCCCAATTTGC 176
 Db 148 TTACGGTTCATTTGCAAGGGTCTGATTAAGTTGAAGTGGGTGAAAGTAAACAAATTTAA 207
 QY 177 AAACGAGAAATGGGAAATATACATATCATCAATACAG--AGAACCCAGAGAAATTGA 233
 Db 208 AGATGGGAGTGGCTGTACCATATTAAGTCTTAAAGTGGCTTCATTCCTGTTTCATAGG 267
 QY 234 GTACGCAATTAATTCATTTTATGAGCAACTATTTATTCATGCTTTATGATTTATTCACCC 293

Db 268 CGGATTTAAGAACCCCTTTTCAATTGAAAAAGTTATCGCTACGATATTAGCTACTTTAAACC 32.27

QY 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTTACTCAGACCCTGGATATCA 34.44

Db 328 TAACATGAGCAGACTACCTGCATATGAAACTTGTTGCTTATGATATATTTTCTCTGATGATGC 38.87

QY 345 TTCAACAAGAAAGATCTGAAAACCAAGACATCCTCGAATGAGAAAAACATTTCTTTACCA 40.04

Db 388 CTACCACTTCTCAGAGAGATAGCCTTACCAACAATCTGTGACAACAAGAAATGAGTTTCA 44.77

QY 405 TTCTGTGCCATTACCGAAGTGGAAAAAGACCGGATTAGTTCAATCGGCAGATTAAGTGC 46.44

Db 448 TTTCGACACAAATTTGAAGAAAGTTCCCAAAAACAGGGCTGGGGCTCTCCGGCAGGTTTAGTCA 50.77

QY 465 AGTTGTGCGACAAGTTATTATATCCATTATCCC-----CAGTTTATCAGTAC 51.55

Db 508 AGTTTTAATCAAGCTTTGGCCTCTCTTTTGTATGAGACCTGGAAAAATATATGACAA 56.77

QY 516 GAATTAAGATTTTTTGGACAAACGTTGACAGATTTGACATTTGTTATGCCCAAAAAAGAT 57.55

Db 568 ATATAGAGAGATTTATCATATATTTTACACAAGTTGCTCATTTGCAAGCTCAGGGTAAAT 62.77

QY 576 AGGATCTGGGTTTGATGTTGCAACCTGCATTTATATGCTGATGTTATATGAAAGATTTCA 63.55

Db 628 TGGAAAGCGGTTTGATGATGTAACCGCGACAGCATATGGAATCTATAGATATATGAAAGATTTCC 68.87

QY 636 GCCAGCTTTGATPAATGACGCTGTTTCAGGTTCTGAAAGATGATCTGAGAAAGTTCCCAAC 69.55

Db 688 ACCCCCATTAATCTCTAATTTGCCAGATATTT-----GGAAGTGCTACTTACCGCAGTAA 74.11

QY 696 AGAGTTGAAAAAATGATTGAAAGATACCTGGGAATTCAAAACATGAAAGATGATCATTAAC 75.55

Db 742 ACTGGCCGATTTGGTTGATGAAAGAAAGACTGGAAATATTAAGATTAAGAAAGTAAACCATTTACC 80.11

QY 756 ATACGAATCAAATTAATATGAGTGACGTCAAAGGCTGCTCAGAAACACCCAAATTTGGT 81.55

Db 802 TTTCGGATTTAATCTTTATGATGATGGCGATTTTAAAGATGCTTCAGAAACAGTAAATCTGGT 86.11

QY 816 ATCAGAGATATCCCATGAGAAAAAGAAAAAGCAGAAAGAAAGCTGCTGTGTGATATACCA 87.55

Db 862 CCAAGAGGTAAAAAATTTGATATGATTTCCGATATCCAGAAAGCTTGAAAAATATATACGA 92.11

QY 876 GCTTATATGCGCAATTTACAGTTTATGAAAGAAAT 91.11

Db 922 ACTCGATCATGCAAAATTCGATTTATGATGATGAGACT 95.77

XX	RESULT 22
XX	ID AAD31027
XX	AD31027 standard; DNA; 8224 BP.
XX	AC AAD31027;
XX	DT 29-AUG-2003 (revised)
XX	DT 31-MAY-2002 (first entry)
XX	DE
XX	Operon E DNA encoding mevalonate pathway and IPP isomerase.
XX	Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
XX	isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
XX	transgenic plant; Yeast; phosphomevalonate kinase; HMGRt; HMGs; AACT;
XX	mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MDD;
XX	acetosuccinyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
XX	PMK; HMG-CoA reductase; IPP isomerase; IPPi; ds.
XX	Saccharomyces cerevisiae.
OS	Arabidopsis thaliana.
OS	Streptomyces sp.
OS	Rhodospirillum rubrum.
OS	Chimeric.
XX	MO200210398-A2.

XX 07-FEB-2002.
PD
XX
XX 31-JUL-2001; 2001WO-US024037.
PF
XX
XX 31-JUL-2000; 2000US-0221703P.
PR
XX
PA (HAHN/) HAHN F M.
RA (KUEH/) KUEHNLE A R.
XX
XX Hahn FM, Kuehnle AR;
PI
XX
DR WPI; 2002-217122/27.
XX
XX
PT Use of specific genes of mevalonate and isoprenoid biosynthetic pathways,
PT for providing a cell with hebbicide or antibiotic resistance, and for
PT providing transformed cells having increased isoprenoid production.
XX
PS Claim 77; Page 145-149; 193pp; English.

The invention relates to the use of specific genes of the mevalonate and isoprenoid biosynthetic pathways and inactive gene sites (pseudogenes). Genes of the invention are used to enhance biosynthesis of isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid pathway derived products in the plastids of transgenic plants and microalgae, for producing herbicide or antibiotic resistant transgenic plants and microalgae, for providing transformed cells with increased isoprenoid production compared to non-transformed cells, and for providing a cell with an inserted polynucleotide sequence encoding one or more products of interest. The present sequence is operon B DNA encoding the entire mevalonate pathway. This operon contains *S. cerevisiae* orf8 encoding phosphomevalonate kinase (PMK), mevalonate kinase (MKV), and *A. thaliana* orf8 encoding 3-hydroxy-3-methylglutaryl-thiolase (AACT) and *A. thaliana* orf9 encoding 3-hydroxy-3-methylglutaryl-coenzyme A (HMGCoA) synthase (HMGSS), *Streptomyces* sp. CL1501 orf encoding HMG-CoA reductase (HMGR) and *R. capsulatus* DNA encoding IPP isomerase (IPPI). (Updated on 29-AUG-2003 to standardise OS field)

Sequence 8224 BP; 2228 A; 1847 C; 2016 G; 2133 T; 0 U; 0 Other;

Query Match	10.2%	Score 132;	DB 6;	Length 8224;
Best Local Similarity	50.5%;	Pred. No. 1.7e-24;		

3 GTCAAAAGCATTTATAGTCAGCCGTGGAAAAAGCATTTCTTGCTGGATTTTGGTCTTGA 62
 43 GTTAGAGCCCTTCACTGCCCCAGGAAAGCGTTACTAGCTGTGGATTTTAGTTTGTGA 102
 63 GCCAATTTATGATGCTTAATGTGA CAGCATTTGTCA GGAATGCAATGCAATTAATCAAC 122
 103 TACAAAATATGAAGATTTGTATGTCGATTTATCGGCAAGAAATGCATGCTGTAAGCCCATCC 162
 123 AAAAGAAAC-----CAGTTGAAAAGATCTAGAATCAAAATTTCTTCA CCCCATTGTC 176
 163 TTACGGTTCAATTCGAAAGGCTCTGATTAAGTTTGAAGTGGTGTGAAGAAATGAATCAATTTAA 222
 177 AAAGGGAATGGGAATATCA CATATATCAATACAG--AGAA CCGCAAGAA GTTCA 233
 223 AGATGGGAGTGGCTGTACATATAAGCTAAAGGCGTTCATTCCTGTTTGCATAGG 282
 234 GTCAAGCATTAATCCATTTTAAAGGCAACTATATTCATGCTTTAGTTATATTCAC 293
 283 CGGATCTAAGAACCTTTCATTTGAAAAAGTATGCTAATCCATATTAGCTACTTTAAACC 342
 294 GACGAAAGCATTTGATCTTGAAA-----TCATATTATTCAGACCTCGGATATCA 344
 343 TTAATATGACGACTATCTGCAATTA AAAA CTGTTGTTATATGATATTTTCTCTGATGATGC 402
 345 TTCAAGAAGAGTACTGAAACCAAGACATCTCGAATGGAGAAAAAATTTCTTTTACCA 404
 403 CTACCATTTCTCAGAGAGATAGCGTTTACCGAACATCGTGGCAACAGAAATTTGAAGTTTCA 462
 405 TTCTCGGCATTTACCGAATGGA AAAAGACCGGATATAGTTTCATCTGGCAGGATTAAGTTC 464


```

Db      463 TTCGACAGAAATGGAAGAGTTCCCAAAACAGGCGCTGGCTCCTCGGACGGTTTATGTCAC 522
Qy      465 AGTTGTGGCAGCAAGTTTATTTATCCATTATTC-----CAATGTTATCAGTAC 515
Db      523 AGTTTAACTACAGCTTTGGCCTCTTTTGTATGACCTGGAAATATATGTAGACA 582
Qy      516 GAATTAAGATATTTTGCACACAGTTGACAGATTGTCATTTTATGCCCAAAAAAGAT 575
Db      583 ATATAGAGAAAGTATTCATATATTTAGCAAGTGTCTATTTGCAAGCTCAGGGTAAAT 642
Qy      576 AGGATCTGGGTTTGTATTTGCAACTGCAATTATGCTGTGATTTATATAGAAATTCA 635
Db      643 TGGAAAGCGGTTGTATGTAGCGCGGACACATATGATCTATCAGATATAGAAAGATTCCC 702
Qy      636 GCCACCTTGAATAATGAGCTTTTCAGGTTCTAGAAATGATCCGAAAGTTCCCAAC 695
Db      703 ACCCGCATTAATCTCTAATTTTGCAGATATTT-----GGAAGGCTACTTACCGCAGTAA 756
Qy      696 AGAGTTGAAAAAATTGATTTGAAAGTAACTGGGAATTCAACATGAAAGATGTACATTACC 755
Db      757 ACTGGCGCATTTGGTTGATGAGGAAGACCTGGAATTTAGATTAAAGTAAACATTAC 816
Qy      756 ATACGAATCAAGTTATTTATGAGTGACGTCAAGGGTGCTCAGAAACCCCAATTGCT 815
Db      817 TTCGGGATTAATCTTATATGATGAGCGCATATTAAGATGGTTGAGAAACAGTAAACTGGT 876
Qy      816 ATCAGAGATCTCCCAATGGAAGAAAGAAAGCCAGAAAGAACTGCTGTGTATATACCA 875
Db      877 CCAGAAAGTAAATAATTTGATGATTCGATATGCCAGAAAGCTTGAAATATATACAGA 936
Qy      876 GCTTAATAGTCCCAATTTTACGTTTATGAAGAAAT 911
Db      937 ACTGCATCATGCAATTTCTAGATTTATGATGACT 972

RESULT 23
AADD1025
ID      AADD1025 standard; DNA; 8235 BP.
XX
XX      AADD1025;
AC
XX      29-AUG-2003 (revised)
DT
XX      31-MAY-2002 (first entry)
XX
XX      Operon C encoding mevalonate pathway and IPP isomerase.
XX
XX      Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
XX      isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
XX      transgenic plant; yeast; phosphomevalonate kinase; HMGR; AAC;
XX      mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MDP;
XX      acetoacetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
XX      MVK; PMK; HMG-CoA reductase; IPP isomerase; IPPi; db.
XX
XX      Saccharomyces cerevisiae.
OS      Arabidopsis thaliana.
OS      Rhodospirillum rubrum.
OS      Chimeric.
XX
XX      WO200210398-A2.
XX
XX      07-FEB-2002.
XX
XX      31-JUL-2001; 2001WO-US024037.
XX
XX      31-JUL-2000; 2000US-0221703P.
XX
XX      (HAIN/) HAIN F M.
XX      (KUEH/) KUEHNLE A R.
XX      Hahn FM, Kuehnle AR;
XX      MPI, 2002-217122/27.

```

```

XX      XX
PT      Use of specific genes of mevalonate and isoprenoid biosynthetic pathways,
PT      for providing a cell with herbicide or antibiotic resistance, and for
PT      providing transformed cells having increased isoprenoid production.
XX      Claim 77, Page 136-140; 193pp; English.
XX
XX      The invention relates to the use of specific genes of the mevalonate and
XX      isoprenoid biosynthetic pathways and inactive gene sites (pseudogenes).
XX      Genes of the invention are used to enhance biosynthesis of isopentenyl
XX      diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid
XX      pathway derived products in the plastids of transgenic plants and
XX      microalgae, for producing herbicide or antibiotic resistant transgenic
XX      plants and microalgae, for providing transformed cells with increased
XX      isoprenoid production compared to non-transformed cells, and for
XX      providing a cell with an inserted polynucleotide sequence encoding one or
XX      more products of interest. The present sequence is operon C DNA encoding
XX      the entire mevalonate pathway. This operon contains S. cerevisiae oris
XX      encoding phosphomevalonate kinase (PMK), mevalonate kinase (AAC),
XX      mevalonate diphosphate decarboxylase (MDP), acetoacetyl thiolase (AACT);
XX      A. thaliana oris encoding 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-
XX      CoA) synthase (HMGS), HMG-CoA reductase (HMGR) and R. capsulatus oris
XX      encoding IPP isomerase (IPPI). (Updated on 29-AUG-2003 to standardise OS
XX      field)
XX
XX      Sequence 8235 BP; 2314 A; 1679 C; 1971 G; 2271 T; 0 U; 0 Other;
XX
XX      Query Match      10.2%; Score 132; DB 6; Length 8235;
XX      Best Local Similarity 50.5%; Pred. No. 1.7e-24;
XX      Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
Qy      3 GTCAAAAGCATTTAGTACGACCTGAAAGAAAGCATTTCTGTGTGTGATATTTGTTCTGA 62
Db      28 GTTGAGAGCTTCAAGTGGCCCAAGGAAAGCTTACTAGCTGTGATATTTAGTTTAA 87
Qy      63 GCCAATTTATGATGCTTATGACAGCATTTGATCAAGATGCAATGCAATGATTAACACC 122
Db      88 TACAATAATATGAAGCATTTGTAGTCGCAAGAAATGCAATGCTGAGCCATCC 147
Qy      123 AAAAGAAC-----CAGTTTGAAGAAATCTAGATCAAAATTTCTTCAACCCCAATTG 176
Db      148 TTACGGTTCATTTGCAAGGCTGTATGATTAAGTGAAGCTGTGAAAGTAAACAATTAA 207
Qy      177 AAACGAGATGGAATATCATATCATCAATCAAG---AGAAAGCCAGAAAGTTCA 233
Db      208 AATGAGGAGTGGCTGTACATATTAAGTCTTAAAGTGGCTTATTCCTGTTCCATAG 267
Qy      234 GTACGCAATTAATTCATTTTGAAGCACTATATTCATGTTTGAAGTTTCAATCC 293
Db      268 CGGATCTAAGAACCTTTTATGAAAGTATGCTAAAGTATTTAGTACTTTAAAC 327
Qy      294 GACCGAAGATTTGATCTTGA-----TCATCATTTTCTAGACCTCGATATCA 344
Db      328 TTAACATGACGACTCTGCAATGAAGAACTGTTCGTATTTGATTTTCTGATGATGC 387
Qy      345 TTCACAGAAGATGATGAAACCAAGACATCTCGAATGGAAGAAACAATTTCTTACCA 404
Db      388 CTACATTTCTCAGAGATGAGCTTACCAACATGCTGCAACAGAAAGTTAGTTTCA 447
Qy      405 TTCTGTGTCATTTACGAAAGTGAAGAACCCGATTTAGTTTCTGCGAGATTAAGTGC 464
Db      448 TTCGACAGAAATTTGAAGAAAGTTCCCAAAACAGGCGTGGCTCCTCGGACGTTTGTGC 507
Qy      465 AGTTGTGGCAGCAAGTTTATTTATCCATTATTC-----CAATGTTATCAGTAC 515
Db      508 AGTTTAACTACAGCTTTGGCCTCTTTTGTATGACCTGGAATAATATGTAGACA 567
Qy      516 GAATTAAGATATTTTGCACACAGTTGACAGATTGTCATTTTATGCCCAAAAAAGAT 575
Db      568 ATATAGAGAAAGTATTCATATTTTAGCAAGTGTCTATTTCAAGCTCAGGGTAAAT 627
Qy      576 AGGATCTGGGTTGTATTTGCAACTGCAATTATGCTGTGATTTATATAGAAATTCA 635

```

Db 628 TGAAGCGGGTTGATGTAGCGGCGGCGACATATGATCTATGATATAGAAATTTCC 687
Qy 636 GCCAGCTTTGATTAATGACGGTGTTCAGTTCTAGAAAGNATCCGAGAAAGTTCCCCAC 695
Db 688 ACCCGCATTAATCTCTAATTTGCCAGATATTT-----GGAAGTGTCTTACGCGAGTAA 741
Qy 696 AGAGTTGAAAAAATGATTTGAAAAGTAACTGGGAATTCMAACATGAAAGATGATCAATTACC 755
Db 742 ACTGGCGCATTTGGTGTGATGAGAAAGACTGGAATATTAAGATTAAAGTAACCATTTACC 801
Qy 756 ATACGGAATCAAGTTTATTAATGGTGACGTCAGAGGGTGTCTGAAACACCCAAATTTGGT 815
Db 802 TTGGGATTAATCTTAATGATGGGCGCATATTAAGAAATGTTTCAGAAACAGTAAACCTGGT 861
Qy 816 ATACGAGTACTCCAAATGAAAAAGAAAAAGCCAGAAAGAAAGCTCTGTTGTGTATGACCA 875
Db 862 CCAGAAAGTAAAAAATTTGGTATGATTCGCAATATGCCAGAAAGCTTAAAAATATATACAA 921
Qy 876 GCTTAATAGTCCAAATTTTACAGTTTATGAAGAATT 911
Db 922 ACTGATCATGCAAAATTTAGATTATGATGAGACT 957

RESULT 24

AAD31029

ID AAD31029 standard; DNA; 8400 BP.

XX AAD31029;

AC AAD31029;

DT 29-AUG-2003 (revised)

DT 31-MAY-2002 (first entry)

DE Operon G DNA encoding mevalonate pathway and IDII (IPP isomerase).

XX Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;

KM isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;

KM transgenic plant; Yeast; IPP isomerase; IDII, da.

XX Saccharomyces cerevisiae.

OS Arabidopsis thaliana.

OS Schizosaccharomyces pombe.

OS Chimeric.

PN WO200210398-A2.

XX 07-FEB-2002.

XX 31-JUL-2001; 2001WO-US024037.

XX 31-JUL-2000; 2000US-0221703P.

PA (HANN/) HANN F M.

PA (KUEH/) KUEHNLE A R.

PI Hahn FM, Kuehnle AR;

XX MPI; 2002-217122/27.

PT Use of specific genes of mevalonate and isoprenoid biosynthetic pathways,

PT for providing a cell with herbicide or antibiotic resistance, and for

PT providing transformed cells having increased isoprenoid production.

PS Claim 77; Page 154-158; 193pp; English.

CC The invention relates to the use of specific genes of the mevalonate and
CC isoprenoid biosynthetic pathways and inactive gene sites (pseudogene).
CC Genes of the invention are used to enhance biosynthesis of isopentenyl
CC diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid
CC pathway derived products in the plasmids of transgenic plants and
CC microalgae, for producing herbicide or antibiotic resistant transgenic
CC plants and microalgae, for providing transformed cells with increased
CC isoprenoid production compared to non-transformed cells, and for
CC providing a cell with an inserted polynucleotide sequence encoding one or

CC more products of interest. The present sequence is operon G encoding
CC entire mevalonate pathway and IDII (IPP isomerase). This operon contains
CC S. cerevisiae, A. thaliana and S. pombe DNA. (updated on 29-AUG-2003 to
CC standardise OS field)

XX Sequence 8400 BP; 2437 A; 1644 C; 1941 G; 2378 T; 0 U; 0 Other;

Qy Query Match 10.2%; Score 132; DB 6; Length 8400;

Best Local Similarity 50.5%; Pred. No. 1.7e-24;

Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

Qy 3 GTCAAAAGCATTTAGTGTGACCTGGAAGAAAGCATTTCTTGTGGATTTGTTCTTGA 62
Db 28 GTTGAAGACCTTCAAGTGTGCGCCAGGAAAGCCTTACTGCTGTGTATTTAGTTTGA 87
Qy 63 GCCAATTTATGATGCTTATGTGACAGATTGTCAACGAATGACGATGATTAACACC 122
Db 88 TACAAAATATGAAAGATTTGTAGTCGATATTCGCAAGAAATGATGCTGTAGCCCATCC 147
Qy 123 AAAAGGAAAC-----CAGTTGAAAGAAATCTAGAAATCAAAATTTCTTCAACCCCATTTGC 176
Db 148 TTACGCTTCATTTGCAAGGCTCTGATTAAGTTGAAGTGTGTGAAAAGTAAACATTTAA 207
Qy 177 AAACGGAATGGGAATATATCATATATCAATATACAG---AGAAGCCAGAGAAAGTTCA 233
Db 208 AGATGGGAGTGGCTGTACATATATAGTCTTAAAGAGGCTTCAATCTGTTTGATAGG 267
Qy 234 GTCAAGCATTAATTCATTTTATGAGGCAATATATTCATGCTTTAGCTTATATTCACC 293
Db 268 CGGATCTAAGAAACCTTTCATTTGAAAAAGTATGCTTACGATTTAGTACTTTAAACC 327
Qy 294 GACGAAGCATTTGATCTTGAAA-----TCATCATTTTACTCAACCCCTGTGATATCA 344
Db 328 TAAATGAGACGATCTACGCAATATGAAACTTGTCTTATTTGATATTTCTCTGATGATGC 387
Qy 345 TTCAAGAAAGATATCTGAAACCAAGACATCTCGAATGGAAGAAAAATTTCTTTTACA 404
Db 388 CTACCATTTCTCAGAGATAGCTGTACCGAAATCGTGGCAAGAAAGATTTGATTTCA 447
Qy 405 TTCTCGGCTCATTTACGAAGTGAAGAAAGACCGATTTGTTTCATCGGCAGATTAAGTGC 464
Db 448 TTGCAAGAAATTAAGAAAGTTCCCAAAACAGGGCTGGGCTCTCGGCAAGTTTAAAGCAC 507
Qy 465 AGTGTGTCACAAAGTTTATTTATCCATTTTATCC-----CAATGTTATCACTAC 515
Db 508 AGTTTAACTACAGCTTTTGGCTCTTTTGTATGAGACTGGAATAATATGAGCAA 567
Qy 516 GAATTAAGATATTTTGCACAAAGTTGCAAGATTGCAATTTGTATGCCAAAAAAGAT 575
Db 568 ATATAGAGAAAGTTATTCATATTTAGCACAAGTTGCTCATTTGTCAGAGGTAAAT 627
Qy 576 AGATCTGGGTTGATGTGCAACCTGCAATTTTATGCTGATTTATGATGAAGATTTCA 635
Db 628 TGAAGCGGGTTGATGTAGCGGGCGGCGACATATGATCTTATGATGAAGATTTCCC 687
Qy 636 GCCAGCTTTGATTAATGACGTTGTTTCAAGTTCTTGAAGAGTATCTGAGAAAGTTCCAC 695
Db 688 ACCCGCATTAATCTCTAATTTGCCAGATATTT-----GGAAGTGTCTTACGCGAGTAA 741
Qy 696 AGAGTTGAAAAAATGATTTGAAAAGTAACTGGGAATTCMAACATGAAAGATGATCAATTACC 755
Db 742 ACTGGCGCATTTGGTGTGATGAGAAAGACTGGAATATTAAGATTAAAGTAACCATTTACC 801
Qy 756 ATACGGAATCAAGTTTATTAATGGTGACGTCAGAGGGTGTCTGAAACACCCAAATTTGGT 815
Db 802 TTGGGATTAATCTTAATGATGGGCGCATATTAAGAAATGTTTCAGAAACAGTAAACCTGGT 861
Qy 816 ATACAGATATCTCAATGAAAAAGAAAAAGCCAGAAAGAAAGCTCTGTTGTGTATGACCA 875
Db 862 CCAGAAAGTAAAAAATTTGGTATGATTCGCAATATGCCAGAAAGCTTAAAAATATATACAA 921
Qy 876 GCTTAATAGTCCAAATTTTACAGTTTATGAAGAATT 911

DB	9222	ACTGCATCATGCAATTCAGATTATGATGACT	957
RESULT 25			
ADP23289			
ID	ADP23289	standard, DNA; 9253 BP.	
XX			
XX			
AC	ADP23289;		
XX			
DT	12-FEB-2004	(first entry)	
XX			
DE	Isopentenyl pyrophosphate synthetase genes operon.		
XX			
KM	isopentenyl pyrophosphate; mevalonate pathway;		
KM	isopentenyl pyrophosphate synthetase; isoprenoid; pharmaceutical;		
KW	nutritional; flavouring agent; agricultural pest control agent;		
KW	acetoacetyl-CoA thiolase; atob; isopentenyl pyrophosphate synthetase gene;		
KX	gene; ds.		
OS	Escherichia coli.		
XX			
PN	US2003148479-A1.		
XX			
PD	07-AUG-2003.		
XX			
PE	06-DEC-2001; 2001US-00006909.		
XX			
PR	06-DEC-2001; 2001US-00006909.		
XX			
PA	(KEAS// KEASLING J.		
PA	(MART// MARTIN V.		
PA	(PITE// PITERA D.		
PA	(KIMS// KIM S.		
PA	(WITH// WITHERS S T.		
PA	(YOSH// YOSHIKUNI Y.		
PA	(NEWM// NEWMAN J.		
PA	(KILE// KHELEBNIKOV A V.		
PI	Keasling J, Martin V, Pitera D, Kim S, WITHERS ST, Yoshikuni Y;		
PI	Newman J, Khelebnikov AV;		
DR	WPI; 2004-020454/02.		
XX			
PT	Synthesizing isopentenyl pyrophosphate in a host microorganism, useful		
PT	for pharmaceutical purposes, comprises introducing into the microorganism		
PT	heterologous nucleic acid sequences coding for an enzyme in the		
PT	mevalonate pathway.		
XX			
PS	Claim 45; SEQ ID NO 7; 40pp; English.		
XX			
CC	The invention describes a method of synthesizing isopentenyl		
CC	pyrophosphate (IP) in a host microorganism. The method comprises		
CC	introducing into the host microorganism a plurality of heterologous		
CC	nucleic acid sequences each coding for a different enzyme in the		
CC	mevalonate pathway for producing isopentenyl pyrophosphate. The method is		
CC	useful in synthesizing isopentenyl pyrophosphate and the derived		
CC	isoprenoids. The isoprenoids may be used in pharmaceuticals,		
CC	nutraceuticals, flavouring agents or in agricultural pest control agents.		
CC	This sequence represents an operon encoding 6 genes of responsible for		
CC	isopentenyl pyrophosphate synthesis.		
XX			
XX			
Sequence	9253 BP; 2690 A; 1912 C; 2150 G; 2501 T; 0 U; 0 Other;		

	Query Match	Best Local Similarity	Score 132;	DB 12;	Length 9253;
	Matches 473;	Conservative	0;	Mismatches 430;	Indels 33; Gaps 5;
Qy	3	GTCAAAACCATTTAGTCGACCTGTGAAGCAATTTCTGTGTGTGATATTATTTGTTCTTGA	62		
Db	6259	GTTAGACCTTCATGCGCCCGAGGAAACGTTACTAGCTGTGTGATATTATTTAGTTTAA	631		
Qy	63	GCCAAATTTATGATGCTTATGTGACAGCATTTGTCATCAGCAATGATGACAGTTAATAACCC	122		

[illegible]

XX Saccharomyces cerevisiae.
OS Escherichia coli.
OS Synthetic.
XX
XX US2004005678-A1.
XX
XX 08-JAN-2004.
XX
XX 09-APR-2003; 2003US-00411066.
XX
XX 06-DEC-2001; 2001US-00006909.
XX
XX (KEAS/) KEASLING J.
XX (MART/) MARTIN V.
XX (PITER/) PITERA D.
XX (WITH/) WITHERS S T.
XX (NEWM/) NEWMAN J.
XX
XX Keasling J, Martin V, Pitera D, Withers ST, Newman J;
XX WPI: 2004-120864/12.
XX
XX Synthesizing amorpho-4,11-diene in a host cell, useful as
XX pharmaceuticals, comprises introducing nucleic acid sequences, each
XX coding for a different enzyme in the mevalonate pathway for producing
XX isopentenyl pyrophosphate.
XX
XX Disclosure, SEQ ID NO 7; 75pp; English.
XX
XX This invention relates to a novel method of synthesizing amorpho-4,11-
XX diene in a host microorganism which comprises introducing into the host
XX microorganism heterologous nucleic acid sequences, each coding for a
XX different enzyme in the mevalonate pathway for producing isopentenyl
XX pyrophosphate, and introducing into the host microorganism a DNA fragment
XX coding for an optimised synthetic amorpho-4,11-diene synthase gene. The
XX nucleic acid sequences, DNA fragments, vectors and host cells are useful
XX for synthesizing isopentenyl pyrophosphate, amorpho-4,11-diene and epi-
XX cedrol used as pharmaceuticals, nutraceuticals, flavouring agents, and epi-
XX agricultural pest control agents. The present sequence is that of an
XX operon which may be used in the method of the invention.

XX
XX Sequence 9253 BP; 2690 A; 1912 C; 2150 G; 2501 T; 0 U; 0 Other;

Query Match 10.2%; Score 132; DB 12; Length 9253;
Best Local Similarity 50.5%; Pred. No. 1.8e-24;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

XX
XX 3 GTCAAAAGCATTGACCTGGAAGAAACATTTCTGCTGGTGAATATTTGGTTCTTGA 62
XX 6259 GTTGAAGGCTTCAGTCCAGGAAAGCGTTACTAGCTGGTGAATATTTAGTTTGA 6318
XX
XX 63 GCCAATTATGATGCTTATGAGCAAGCATGTCATCAAGATGATGATGATTAACACC 122
XX DB 6319 TCAAAATATGAGCAATTTGATGCTGATGATGCAAGATGATGCTGATGATGCA 6378
XX 123 AAAAGAAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCAACCCCAATTGC 176
XX DB 6379 TTACGTTTCATTCGAAGGCTCTGATAGTTGAAGTGGTGTGAAAAGTAAACAATTTAA 6438
XX 177 AAACGAGATGGAATATCATATCATATCAATACAG--AGAACCCAGAGAGATTTCA 233
XX DB 6439 AGATGGGAGTGGCTGACCATATAGTCTTAAAGTGGCTTCAATCTGTTGATAGG 6498
XX 234 GTACGATTAATATCATTTTATAGAGCAATATATTCATGTTAGCTTATATCAACC 293
XX DB 6499 CGGATCTAAGAACCTTTTATGAAAAAGTTATGCTAAGCTATTAAGCTTTAAACC 6558
XX 294 GACGAAACATTTGATCTTGA---TATCATTTTATCTAGACCTCTGGATATCA 344
XX DB 6559 TAAAGTGAAGCACTAGTCAATAGAACTTGTGTTATGATATTTTCTGATGATGTC 6618
XX 345 TTCACAGAAAGATCTAGAAACCAAGACATCTCTGAATGAGAAAAACATTTTCTTACCA 404

DB 6619 CTACCATTCCTCAGAGAGATAGGCTTACCGAATCTGTGGCAACAGATGATGTTTCA 6678
XX
XX 405 TTCTGTCGATTTACCGAAGTGAAGAAAGCCGATTTAGTTTCATGCGAGATTAAGTGC 464
XX DB 6679 TTGCAAGAAATTAAGAAAGTTCACCAAAACAGGGCTGGGCTCTCGCAGGTTTACGAC 6738
XX 465 AGTTGTCACAAAGTTATTTATCCCATTTTATCC-----CATGTTATCAGTAC 515
XX DB 6739 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATGCACTCGAATAATTAATGACAA 6798
XX 516 GAATTAAGATATTTTGCACAAAGTTGCAACAGATGCACTTTATGCCCCAAAAAGAT 575
XX DB 6799 ATATGAGAAAGTTATTCATTAATTTAGCAAGTTGCTCATGTCAGGCTCAAGGTTAAAT 6858
XX 576 AGCATCTGGGTTGATGTTGCAACGCAATTTATGCTGATTTATATAGAAATTTCA 635
XX DB 6859 TGAAGCGGTTTATGATGAGCGCGGCGGCAATATGATCTATGATGATGAAAGATTCCC 6918
XX 636 GCCAGCTTTGATTAATGACGTTTCAAGTTCTAGAAAGTATCTGAGAAAGTTCCCCAC 695
XX DB 6919 ACCCGCATTAATCTCTAATTTGCCAGATAT-----GGAAGTCTACTTAACGCACTAA 6972
XX 696 AGAGTTGAAAAAATTTGATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 755
XX DB 6973 ACTGCGCATTTGGTTGATGAGAAAGCTGAAATATTAACATTAACCATTTACC 7032
XX 756 ATACGGAATCAAGTTATTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 815
XX DB 7033 TTCCGGATTAATCTTATGATAGGCGCATTTAAGATGTTTCAAGAAACGATTAACCTGGT 7092
XX 816 ATACGAGTACTCAATGAGAAAGAAAGAAAGCAAGAAAGCTCTGTTGATGACCA 875
XX DB 7093 CCAAGAGTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7152
XX 876 GCTTAATAGTCCAAATTTTACAGTTTATGAAAGAT 911
XX DB 7153 ACTGATCATCAAAATTTCTAATTTATGATGATGATGATGATGATGATGATGATGATGAT 7188

RESULT 27

AAD31037 standard; DNA; 13917 BP.

AAD31037;

29-AUG-2003 (revised)
31-MAY-2002 (first entry)

Plastid transformation vector pK04 DNA.

Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
transgenic plant; Yeast; phosphomevalonate kinase; HMGs; ACT;
mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MD;
acetylacetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
HMG; PMK; HMG-CoA reductase; ds.

Saccharomyces cerevisiae.
Arabidopsis thaliana.
Chimeric.

MO200210398-A2.

07-FEB-2002.

31-JUL-2001; 2001MO-US024037.

31-JUL-2000; 2000US-0221703P.

(HAHN/) HAHN F M.
(KUEH/) KUEHNLE A R.

PI Hahn FM, Kuehnle AR;

XX WPI: 2002-217122/27.

XX Use of specific genes of mevalonate and isoprenoid biosynthetic pathways,
PT for providing a cell with herbicide or antibiotic resistance, and for
PT providing transformed cells having increased isoprenoid production.

PS Claim 77; Page 161-168; 193pp; English.

XX The invention relates to the use of specific genes of the mevalonate and
CC isoprenoid biosynthetic pathways and inactive gene sites (pseudogenes).
CC Genes of the invention are used to enhance biosynthesis of isopentenyl
CC diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid
CC pathway derived products in the plastids of transgenic plants and
CC microalgae, for producing herbicide or antibiotic resistant transgenic
CC plants and microalgae, for providing transformed cells with increased
CC isoprenoid production compared to non-transformed cells, and for
CC providing a cell with an inserted polynucleotide sequence encoding one or
CC more products of interest. The present sequence is a plastid transformation
CC vector pHK04 containing operon B DNA which encodes the entire mevalonate
CC pathway. This operon contains S. cerevisiae orf8 encoding
CC phosphomevalonate kinase (PMK), mevalonate kinase (MVK), mevalonate
CC diphosphate decarboxylase (MDD), acetoacetyl thiolase (AAT) and A.
CC thaliana orf8 encoding 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA)
CC synthase (HMGs) and HMG-CoA reductase (HMGRL). (Updated on 29-AUG-2003 to
CC standardise OS field)

XX Sequence 13917 BP; 3882 A; 2923 C; 3063 G; 4049 T; 0 U; 0 Other;

Query Match 10.2%; Score 132; DB 6; Length 13917;

Best Local Similarity 50.5%; Pred. No. 2.1e-24;

Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTGACCTGGAAGACATTTCTGCTGTGATTTGGTCTTGA 62
DB 4137 GTTGAAGAGCTTCAGGCCCGAGGAAAGCTTACGCTGGATATTGTTTGA 4196
QY 63 GCCAATTTATGATGCTTATGACAGCATTTGTCATCAGAAAGCATGCTATTAACAC 122
DB 4197 TACAAATATGAAAGCTTTGATGTCGATTTGCGCAAAAGCATGCTGAGCCCATC 4256
QY 123 AAAAGAAC-----CAGTTGAAGAAATCTAGAAATCAAAATTTCTTACCCCAATTTGC 176
DB 4257 TTAAGGTTTCATGCAAGGCTCTGATTAAGTTTGAAGCGCTGCGAAAGTAAACATTTTA 4316
QY 177 AAACGAGAGTGGAAATATCATATCATCAATATCAG--AGAACCCAGAGAAAGTTCA 233
DB 4317 AGATGGGGAGTGGCTGTACATATTAAGTCTTAAAGTGCTTCATTCCTGTTTCATAG 4376
QY 234 GTACAGCATTAATCCATTTTAGAGGACATATTCATTCGTTTAACTTATTTCAAC 293
DB 4377 CGGAATCTAAGAACCTTTTCATTTGAAAGTATTCGTTAAGTAAAGTAAAC 4436
QY 294 GACCGAAGCATTGATCTTGA--TCATCATTTTACTCAGACCTGATATCA 344
DB 4437 TAAATGAGAGCATCTACTGCAATTAAGAACTTTCGTTATGATTTTCTCGATGATGC 4496
QY 345 TTCAAGAGAGTACTGAAACCAAGACATCTCGAATGAGAAAGAAATTTCTTTACCA 404
DB 4497 CTACACATTTCTCAGAGAGATAGGTTTACCGAACATCGTGCAACAGAAATGAGTTTGA 4556
QY 405 TTCTGTCGTCATTACCGAAGTGGAAAGACCGGATTTAGTTTATCGGCGAGTATGTC 464
DB 4557 TTTCGACAGAAATGAAAGATTCCTCAAAACAGGCGTGGCTCTCGGAGGTTTATGTCAC 4616
QY 465 AGTTGTCGACAGATTTTATTCATTTTATCCCATTTTATCC-----CAATGTTATCAGTAC 515
DB 4617 AGTTTATTAACAGCTTGGCTCTTTTGTATGAGCATGCGAAATTAATGTAGACA 4676
QY 516 GAATTAAGATTTTTCACACAGTGGACAGATTGCAATGTTATGCTCCCAAAAAAGAT 575
DB 4677 ATATAGAGAAATTTATTAATTTAGACAAAGTGTCTATTTCAAGCTCAGGGTAAAT 4736

QY 576 AGATCTGGGTTGATGTTGCAATGCAATTTATGCTGATTTATAGAGATTGA 635

DB 4737 TGGAAAGCGGTTTATGATGACCGCGACATATGATCTATCAATATGAAGATTCC 4796

QY 636 GCCACCTTGAATTAAGCGTGTTCAGGTTCTAGAAAGTATCTGAGAACTTCCAC 695

DB 4797 ACCCCATTAATCTCTAATTTGCGAGATTT-----GGAAGTGTACTTACGGCAGTAA 4850

QY 696 AGAGTTGAAAAAATGATTTGAAGATTAACGGAATTTCAACATGAAAGTATCACTTACC 755

DB 4851 ACTGCGCATTTGTTGTTGAAGAAAGACTGGAATTTAGATTAAGAACCATTTTACC 4910

QY 756 ATACGAATCAAGTATTATTAATGAGTACGTCAGAGGTGCTCAGAAACCCAAATTTGCT 815

DB 4911 TTTCGGATTAATCTTTATGATGATGCGCATTTTAAGATGTTTCAAGAAACGTAATTTGCT 4970

QY 816 ATCAGAGTACTCCAAATGGAAGAAAGGCAAGAAAGCTCTGTTGTATGACCA 875

DB 4971 CCAGAAAGTAAATAATGTTGATGATTCGCAATGCGCAAGAAAGCTTGAATAATATACAGA 5030

QY 876 GCTTAATAGTCCATTTTACAGTTTATGAGAAAT 911

DB 5031 ACTCGATCATGCAAAATTTTATGATTTATGATGAGTACT 5066

RESULT 28

AAD31039
ID AAD31039 standard; DNA; 14623 BP.

XX AAD31039;

XX 29-AUG-2003 (revised)

DT 31-MAY-2002 (first entry)

DE Plasmid transformation vector pHK08 DNA.

KW Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
KW isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
KW transgenic plant; yeast; IPP isomerase; ID11; db.

XX Saccharomyces cerevisiae.

XX Arabidopsis thaliana.

OS Schizosaccharomyces pombe.

OS Chimeric.

PN WO200210398-A2.

PD 07-FEB-2002.

PP 31-JUL-2001; 2001WO-US024037.

XX 31-JUL-2000; 2000US-0221703P.

PA (HAHN/) HAHN F M.

PI (KUEH/) KUEHNLE A R.

XX Hahn FM, Kuehnle AR;

XX WPI: 2002-217122/27.

PT Use of specific genes of mevalonate and isoprenoid biosynthetic pathways,
PT for providing a cell with herbicide or antibiotic resistance, and for
PT providing transformed cells having increased isoprenoid production.

PS Claim 77; Page 173-181; 193pp; English.

XX The invention relates to the use of specific genes of the mevalonate and
CC isoprenoid biosynthetic pathways and inactive gene sites (pseudogenes).
CC Genes of the invention are used to enhance biosynthesis of isopentenyl
CC diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid
CC pathway derived products in the plastids of transgenic plants and
CC microalgae, for producing herbicide or antibiotic resistant transgenic

CC plants and microalgae, for providing transformed cells with increased
 CC isoprenoid production compared to non-transformed cells, and for
 CC providing a cell with an inserted polynucleotide sequence encoding one or
 CC more products of interest. The present sequence is plasmid transformation
 CC vector pKH06 containing operon G DNA which encodes the entire mevalonate
 CC pathway and IDI1 (IPP isomerase). This operon contains 8. cerevisiae, A.
 CC thaliana and S. pombe DNA. (updated on 29-AUG-2003 to standardise OS
 CC field)

XX Sequence 14623 BP; 4128 A; 3028 C; 3227 G; 4240 T; 0 U; 0 Other;

Query Match 10.2%; Score 132; DB 6; Length 14623;
 Best Local Similarity 50.5%; Pred. No. 2.1e-24;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTGTGTGATTTGGTCTTGA 62
 DB 2621 GTTGAGAGCCTTCAGAGCCCCAGGAAAGCGTTACTAGCTGGGATTTTATGTTTAA 2680
 QY 63 GCCAATTTATGATGCTTATGTGACAGCATTTGCATCAGATGCATGCAATTATACACC 122
 DB 2681 TACAAAATATGAAAGCATTTGTAGTCGATTTATCGCAAGAAATGCATGCTGAGCCATCC 2740
 QY 123 AAAAGGAAAC-----CAGTTGAAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTGCG 176
 DB 2741 TTAGGCTTCAATTCAGAGGCTCTGATTAAGTTGAAGTCGTTGAAAAAGTTAAACAAATTAA 2800
 QY 177 AAAAGGAAATGGGAAATATCATATCATCAATAATACAG--AGAAAGCCAGAGAAAGTTCA 233
 DB 2801 AGATGGGAGTGGCTGTACATATTAAGTCTTAAAGTGGCTTCACTTCTGTTTCCATAGG 2860
 QY 234 GTACAGCATTAATCCATTTTAAAGGCAACTATATTCATGCTTTAGCTTATATTTCAACC 293
 DB 2861 CGGATCTAAGAAACCTTTTCAATGAAAAAGTTACGCTAACGATTTAGCTACTTTAAACC 2920
 QY 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTCTCAGACCCCGATATCA 344
 DB 2921 TAAATGAGAGCACTACTGCATATGAAACTGTTGCTGTTATGATTTTCTCGATGATGC 2980
 QY 345 TTTCACAGAGATGACTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTACCA 404
 DB 2981 CTACCAATCTCAGAGAGTAGGCTTACCGAACATCGTGGCAACAGAAATTAAGTTTCA 3040
 QY 405 TTCTCGTGCATTAACCGAGTGGAAAAAGCCGATTTAGGTTATGCGCGAGATTAGTGTCC 464
 DB 3041 TTTCGACAGAAATGAAAGATTTCCAAAACAGGCGCTCTCGGCGAGTTTATGATCAC 3100
 QY 465 AGTTGTCACAGATTTTATATCCCATTTTATCC-----CAATGTTATCAGTAC 515
 DB 3101 AGTTTAACTACAGCTTGGCCTCTTTTGTATCGAAGCTGGAAATATGTTAGACAA 3160
 QY 516 GAATTAAGATTTTTCACACAGCTGGACAGATTGCAATTTGTTATGCCCAAAAAAAGAT 575
 DB 3161 ATATGAGAGAAATTAATTAATTAATTTAGCAAGTCTCTAATGTCAAGCTCAGCGGTAAT 3220
 QY 576 AGGATCTGGGTTGATGTTGCAACGCAATTTATGCTGATGTTATTAAGAAATTTCA 635
 DB 3221 TGGAAAGCGGTTGATGAGCGGCGCAGCAATATGATCTTACAGATTAAGAAATTTCC 3280
 QY 636 GCCAGCTTTGATTAATGACGTTGTTCAAGTTCTAGAAAGTGAATCTGGAAGTTTCCAC 695
 DB 3281 ACCCGCATTAATCTCTAATTTGCCAGATATTT-----GGAAGTCTACTTACGGCAGTAA 3334
 QY 696 AGAGTTGAAAAAATGATTTGAAAAGTAACTGGGAATTTCAACATGAAAGATGTACATTAC 755
 DB 3335 ACTGGCGGATTTGTTGATGAAAGAACTGGAAATTTACATTAATAAGTAAACCAATTATAC 3394
 QY 756 ATACGGAATCAAGTATTAATGAGGAGCTGAAGGTTGGCTCAAAAACCAATTTGTT 815
 DB 3395 TTTCGGATTAATCTTATGATGAGGCGGATTAATTAAGATGTTTCAAGAAACAGTAAATCGGT 3454
 QY 816 ATCAGAGTACTCCATGAAAAAAGAAAGCAGAAAGAAAGCTCTGTTGTTATGACCA 875

DB 3455 CCAGAAAGTAAAAAATGTTATGATTCGATATGCCAGAAAGCTTGAATAATATACAGA 3514
 QY 876 GCTTAATAGTCCATTTTACAGTTATGAGAAATTT 911
 DB 3515 ACTGATATGATCAAAATTTCAAGATTTATGATGACT 3550

RESULT 29
 AAD31041
 ID AAD31041 standard; DNA; 14623 BP.
 XX
 AC AAD31041;
 XX
 DT 29-AUG-2003 (revised)
 DT 31-MAY-2002 (first entry)
 XX
 DE Plasmid transformation vector pKH06 DNA.

XX Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
 KW isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
 KW transgenic plant; yeast; phosphomevalonate kinase; HMGR; HMGs; AACT;
 KW mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MDD;
 KW acetoacetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
 KW MWK; PKM; HMG-CoA reductase; IPP isomerase; IPPi; ds.

XX Saccharomyces cerevisiae.
 OS Arabidopsis thaliana.
 OS Streptomyces sp.
 OS Rhodobacter capsulatus.
 OS Chimeric.

XX WO200210398-A2.
 XX
 PD 07-FEB-2002.

XX 31-JUL-2001; 2001WO-US024037.
 XX
 PR 31-JUL-2000; 2000US-0221703P.

XX (HAHN/) HAHN F M.
 PA (KUEHN/) KUEHNLE A R.

XX Hahn FM, Kuehnle AR;
 PI
 DR WPI; 2002-217122/27.

PT Use of specific genes of mevalonate and isoprenoid biosynthetic pathways,
 PT for providing a cell with herbicide or antibiotic resistance, and for
 PT providing transformed cells having increased isoprenoid production.

XX Claim 77; Page 185-193; 193pp; English.

XX The invention relates to the use of specific genes of the mevalonate and
 CC isoprenoid biosynthetic pathways and inactive gene sites (pseudogene).
 CC Genes of the invention are used to enhance biosynthesis of isopentenyl
 CC diphosphate (IPP), dimethylallyl diphosphate (DMAPP) and isoprenoid
 CC pathway derived products in the plasmids of transgenic plants and
 CC microalgae, for producing herbicide or antibiotic resistant transgenic
 CC plants and microalgae, for providing transformed cells with increased
 CC isoprenoid production compared to non-transformed cells, and for
 CC providing a cell with an inserted polynucleotide sequence encoding one or
 CC more products of interest. The present sequence is plasmid transformation
 CC vector pKH06 containing operon B DNA which encodes the entire mevalonate
 CC pathway. This operon contains S. cerevisiae orf encoding
 CC phosphomevalonate kinase (PMK), mevalonate kinase (MWK), mevalonate
 CC diphosphate decarboxylase (MDD), acetoacetyl thiolase (AACT) and A.
 CC thaliana orf encoding 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA)
 CC synthase (HMGs); Streptomyces sp. Cl190 orf encoding HMG-CoA reductase
 CC (HMGR) and R. capsulatus DNA encoding IPP isomerase (IPPi). (Updated on
 CC 29-AUG-2003 to standardise OS field)

XX Sequence 14623 BP; 4128 A; 3028 C; 3227 G; 4240 T; 0 U; 0 Other;

Db 1390 CTGATGTCATGCGGACAGTCAGGCTGTGTAGGCGGTGCGTTCCAGATGCGGGGCGC 1449
 QY 1138 TACGATGC 1145
 |||||
 Db 1450 TACGATGC 1457

RESULT 31

ABT19236
 ID ABT19236 standard; DNA; 1356 BP.

AC ABT19236;

DT 16-APR-2003 (first entry)

DE Aspergillus fumigatus essential gene #1594.

KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
 cancer; contamination; biofilm; antibody; immune response; ds.

OS Aspergillus fumigatus.

PN W0200286090-A2.

PD 31-OCT-2002.

PF 23-APR-2002; 2002W0-US013142.

PR 23-APR-2001; 2001US-0285697P.

PR 27-APR-2001; 2001US-0287066P.

PR 05-JUN-2001; 2001US-0295890P.

PR 09-JUL-2001; 2001US-0303899P.

PR 31-AUG-2001; 2001US-0316362P.

XX (ELIT-) ELITRA PHARM INC.

PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

DR WPI; 2003-093124/08.

PT New purified or isolated nucleic acids of essential genes of Aspergillus

PT fumigatus, useful for treating or preventing infections by A. fumigatus,

PT or for treating a non-infectious disease in a subject e.g. cancer.

XX Disclosure; Page; 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as A. fumigatus, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or inhibit formation of an object
 CC by A. fumigatus, or to prevent or inhibit formation of an object
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for
 CC expressing recombinant protein for characterisation, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of A.
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This polynucleotide sequence represents one of the essential
 CC genes of Aspergillus fumigatus of the invention

Sequence 1356 BP; 293 A; 384 C; 384 G; 295 T; 0 U; 0 Other;

Query Match 5.1%; Score 66; DB 8; Length 1356;

Best Local Similarity 45.7%; Pred. No. 4.1e-07;

Matches 547; Conservative 0; Mismatches 595; Indels 54; Gaps 7;

QY 6 AAAAGCATTAGTGCACCTCGGAAAAAGCATTTCTGCTGCTGAGATATTTGTTCTTGAGCC 65
 Db AACGCTCTTGCCTCGGCGCGGCGAAGTCTCTTCACTGCGGGGTTATCTGCTCGACCG 83
 QY 66 AATTATGATGCTTATGTGACGACATGTCATCGAATGATGATGATTAACACCAA 125
 Db CAGCTACACTGGAGATGTGTTCTGCGCTTGAAGCCAGATCCATGCGGAAAGTTTCA 143
 QY 126 AGGAACAGTTGAAAGATCTAGAAATCAAAATTTCTTACCCCAATTTGCAACGAGA 185
 Db TGGAGACAAGAGAAAGCAAGGACAGATGCTGTAACATCCCAAGTTGCGATGGGT 203
 QY 144 TGGAGACAAGAGAAAGCAAGGACAGATGCTGTAACATCCCAAGTTGCGATGGGT 203
 Db ATGGGAATATGATCATATCATCAATACAGAGAACCCAGAGAGTTCACTACGATATAA 245
 QY 186 ATGGGAATATGATCATATCATCAATACAGAGAACCCAGAGAGTTCACTACGATATAA 245
 Db ATGGAGATATGATCATCAAAAGATCGAGAGCGAGCTGACCCATCCAGGCCACGCC--AA 260
 QY 204 ATGGAGATATGATCATCAAAAGATCGAGAGCGAGCTGACCCATCCAGGCCACGCC--AA 260
 Db TCCATTTTGAAGGCAACTATATTCATGTTTGTAGTTATAT-----TCAACGACC 297
 QY 246 TCCATTTTGAAGGCAACTATATTCATGTTTGTAGTTATAT-----TCAACGACC 297
 Db CCGTTTGTGAGACCTCTGAACTTGGCTTGAATCATTAATGATGATGCGGACATC 320
 QY 298 GAAGCATTTGATCTTGAAATATCATTTACTCAGACCTCGATATCATTCACAAAGAT 357
 Db GAAGCATTTGCGGATCATTTATCATTTACTATCTCGCAGACAGATTAATCATTCGAGAC 380
 QY 321 GAAGCATTTGCGGATCATTTATCATTTACTATCTCGCAGACAGATTAATCATTCGAGAC 380
 QY 358 ACTGAAACCAAGACATCTCGAATGAG-----AAAAACAATTTCTTATACATCTCGTGC 413
 Db GGCCTTCTCAAGGCTTCGGAATCTCGGATCTCGGCTGCTAAGAGATTTGTGATCTTGC 440
 QY 414 CATTACGAAGTGAAGAAAGACCGATTAAGTTATCGGACAGATTAAGTCAATGTTGC 473
 Db CTTTCAGAGAGGACACAAAGACAGGCTCGGCTCTTGAAGCGCTTGTGATCCTGCTGCT 500
 QY 474 CAGAACTTATATCCATTTA-----TCCCAATTTATATGATCAAGATTAAGATAT 527
 Db GTGCTCTCTGTATATCCAGCTATGCAACCGACAGATCTGCTCGACGCGCACAA 560
 QY 501 GTGCTCTCTGTATATCCAGCTATGCAACCGACAGATCTGCTCGACGCGCACAA 560
 Db TTTGCAACAGTTGACAGATGACATGTTATGTCCTGCAAAAGATTAAGATTTGAGTT 587
 QY 528 TTTGCAACAGTTGACAGATGACATGTTATGTCCTGCAAAAGATTAAGATTTGAGTT 587
 Db GCTTCAACATCTGCGGACAGGCGGCGGACATCGCTGCGAGGATTAAGTTGAGTTGCT 620
 QY 561 GCTTCAACATCTGCGGACAGGCGGCGGACATCGCTGCGAGGATTAAGTTGAGTTGCT 620
 Db TGAATTTGCACTGCAATTTATGCTGATTTATTAAGATTT-----TCAAGCAGCTT 643
 QY 588 TGAATTTGCACTGCAATTTATGCTGATTTATTAAGATTT-----TCAAGCAGCTT 643
 Db CGATGTTGCACTGCAATTTATGCTGATTTATTAAGATTT-----TCAAGCAGCTT 680
 QY 644 TGAATTAATGACGTGTTCAAGTTCTAGAAAGATCTTGAAGATTTCCCAAGATTA 703
 Db CGAATCAAGTGTGACGCTGTTCTTCAAGCTTCAAGAGCGGTTGTTCCGATCGTAA 740
 QY 704 AAAAATGATTAAGATTA-----ACTGGAATTTCAAAATGAAAGATTAATTAATTA 755
 Db GAGCGCCGACCCCAACATCTCGGAGACCGAATGTCAAGCTTCCGATTAAGTATCC 800
 QY 741 GAGCGCCGACCCCAACATCTCGGAGACCGAATGTCAAGCTTCCGATTAAGTATCC 800
 Db ATGGAATCAATTAATTAATGAGTGAAGTCAAGGCTGCTGAGAAACCAAAATTTGCT 815
 QY 756 ATGGAATCAATTAATTAATGAGTGAAGTCAAGGCTGCTGAGAAACCAAAATTTGCT 815
 Db CCGGGAATCAATTAATGCTCTGAGACGTCAAAATGAGTGTGAGAGATCCATTCATGCT 860
 QY 801 CCGGGAATCAATTAATGCTCTGAGACGTCAAAATGAGTGTGAGAGATCCATTCATGCT 860
 Db ATCAAGATTAATTAATGAGTGAAGTCAAGGCTGCTGAGTGAATTAATTAATTA 875
 QY 816 ATCAAGATTAATTAATGAGTGAAGTCAAGGCTGCTGAGTGAATTAATTAATTA 875
 Db GAGAAAGTTTGTGAGTGCAGAAACAGATTAAGAAAGAGCCATATGCTTGGGCGC 920
 QY 861 GAGAAAGTTTGTGAGTGCAGAAACAGATTAAGAAAGAGCCATATGCTTGGGCGC 920
 Db GCTTAATTAATGCAATTAATTAAGTGAAGTGAAGGATTCGCTGAGAAATTAATTA 935
 QY 876 GCTTAATTAATGCAATTAATTAAGTGAAGTGAAGGATTCGCTGAGAAATTAATTA 935
 Db TCTGCAATGCAACAAAGAGACTTGCCTGGAATCAAGAGCTTGGACAGAGCGGGA 980
 QY 921 TCTGCAATGCAACAAAGAGACTTGCCTGGAATCAAGAGCTTGGACAGAGCGGGA 980
 Db CTCAGACCAAGAGACTTGAATTAAGATTAATTAAGATTAATTTCTGTGAGCTTGA 995
 QY 936 CTCAGACCAAGAGACTTGAATTAAGATTAATTAAGATTAATTTCTGTGAGCTTGA 995
 Db CGAACAATCTCAAGTGAATTAAGAAATGTCGACCTTAATTAATTAAGCTCTG----- 1032

QY 996 GATTAGAAATCATCAGAAAGGGTTACACGATTACACAAATAATCAGAGTTCCATTGA 1055
 DB 1033 -----CGTAACACATCCGTTCCATGACTCAAAAGTCGATGCCAATGCA 1079
 QY 1056 ACCGTATGTCGACACCCAGTTGTGACCGGTGTCTCAAGAGATTCCTGTTGTGTGTG 1115
 DB 1080 GCCGGCGGCTCCAAACCGAGCTACTGACCTCTGTCGAGCTGAGAGGTGTCAATCGTGG 1139
 QY 1116 TGTGTTCCAGGTGCTGTGATATACGATCAATAGCTGATTAGTTGTAATC 1171
 DB 1140 TGTGTTCCAGAGACGAGGGGCTACGACCCATTGTGCTCTCATCCAGACATC 1195

RESULT 32
 AB228940/C
 ID AB228940 standard; DNA; 65 BP.
 XX AB228940;
 AC
 XX 30-JAN-2003 (first entry)
 DT
 XX
 DE Candida gene related tetracycline promoter PCR primer SEQ ID NO 3023.
 XX
 KM Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;
 KM signal transduction; DNA replication; cell division; growth;
 KM proliferation; Candida albicans; fungicide; antifungal; PCR; primer; ss.
 OS
 XX Candida albicans.
 XX
 FN WO200253728-A2.
 PD 11-JUL-2002.
 XX
 PE 26-DEC-2001; 2001WO-US049486.
 XX
 PR 29-DEC-2000; 2000US-0259128P.
 PR 20-FEB-2001; 2001US-00792024.
 PR 22-AUG-2001; 2001US-0314050P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
 DR WPI; 2002-566694/60.
 XX
 PT Constructing strains for identifying gene products as effective targets
 PT for therapeutic intervention, by inactivating in the strain one allele of
 PT a gene and placing other allele of the gene under conditional expression.
 XX
 PS Claim 76; SEQ ID NO 3023; 167pp + Sequence Listing; English.
 XX
 CC The invention relates to constructing (M1) a strain of diploid fungal
 CC cells in which both alleles of a gene are modified, comprising modifying
 CC one allele by insertion or replacement by a cassette having an
 CC expressible selectable marker and modifying other allele by
 CC recombination, of a promoter replacement fragment with a heterologous
 CC promoter, so that expression of the second allele is regulated by the
 CC promoter. (M1) is useful for constructing a strain of diploid fungal
 CC cells in which both alleles of a gene are modified. The diploid fungal
 CC cells having both alleles modified are useful for identifying a gene that
 CC is essential to the survival or growth of a fungus, a gene that
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene
 CC that contributes to the resistance of a diploid fungus to an antifungal
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus
 CC and for identifying a therapeutic agent for treatment of a mammalian
 CC disease. (M1) is useful for identifying a compound which modulates the
 CC activity of a gene product, preferably enzymatic activity, carbon
 CC compound catabolism, biosynthesis, transporter, transcriptional,
 CC translational, signal transduction, DNA replication and cell division
 CC activity. The method is useful for identifying a compound having the
 CC ability to inhibit growth or proliferation of C. albicans cells and for
 CC treating infection by C. albicans. The present sequence is that of a PCR

CC primer used in the method of the invention. Note: The sequence data for
 CC this patent is not represented in the printed specification but is based
 CC on sequence information supplied to Derwent by the European Patent Office
 SQ Sequence 65 BP; 22 A; 16 C; 11 G; 16 T; 0 U; 0 Other;
 Query Match 5.0%; Score 65; DB 6; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.2e-07;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTAAAGCATTTAGTGCACCTGAAAGCATTTCTTGCTGTGATATTGTTCTT 60
 DB 65 ATGTAAAGCATTTAGTGCACCTGAAAGCATTTCTTGCTGTGATATTGTTCTT 6
 QY 61 GAGCC 65
 DB 5 GAGCC 1

RESULT 33
 AAS23667/C
 ID AAS23667 standard; DNA; 90 BP.
 XX AAS23667;
 AC
 XX 04-DEC-2001 (first entry)
 DT
 XX
 DE Tetracycline promoter downstream PCR primer (Tet-Down) #42.
 XX
 KM Gene identification; essential gene; GRACE; pathogenic fungus;
 KM gene replacement and conditional expression; fungal infection;
 KM PCR primer; Tet-Down; tetracycline promoter; ss.
 OS
 XX Escherichia coli.
 XX Candida albicans.
 FN WO200160975-A2.
 PD 23-AUG-2001.
 XX
 PE 20-FEB-2001; 2001WO-US005551.
 XX
 PR 18-FEB-2000; 2000US-0183534P.
 PR
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Roemer T, Jiang B, Boone C, Bussey H;
 DR WPI; 2001-489080/53.
 XX
 PT Identifying genes essential to fungal metabolisms and identifying
 PT potential therapeutic agents that target these genes.
 XX
 PS Disclosure; Page 297; 324pp; English.
 XX
 CC The present invention relates to novel methods for constructing fungal
 CC strains useful for identification and validation of gene products as
 CC targets for therapeutic agents, for creating a collection of identified
 CC essential genes, and screening assays for the discovery of new drugs. The
 CC invention provides the GRACE (gene replacement and conditional
 CC expression) method for the construction of mutant organisms referred to
 CC as GRACE strains of the organism. The invention can be applied to any
 CC organism, particularly a pathogenic fungus e.g. Candida albicans,
 CC Aspergillus fumigatus and Cryptococcus neoformans. The methods are useful
 CC to identify agents that may be used in the treatment of fungal
 CC infections. AAS23626-AAS23686 represent tetracycline promoter downstream
 CC PCR primers (Tet-Down) used in the methods of the present invention
 XX
 SQ Sequence 90 BP; 28 A; 19 C; 18 G; 25 T; 0 U; 0 Other;
 Query Match 5.0%; Score 65; DB 4; Length 90;
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGTCAAAAGCATTTAGTGCACCTGGAAAAAGCATTTCTGTGTGATATTGGTCTT 60
 DB 65 ATGTCAAAAGCATTTAGTGCACCTGGAAAAAGCATTTCTGTGTGATATTGGTCTT 6
 OY 61 GAGCC 65
 DB 5 GAGCC 1
 RESULT 34
 ID ABT21056 standard; DNA; 1455 BP.
 AC ABT21056;
 XX
 DT 16-APR-2003 (first entry)
 DE Aspergillus fumigatus essential gene #3414.
 XX
 DE Aspergillus fumigatus essential gene; Aspergillus fumigatus; infection;
 KM Fungicide; cytosolatic; essential gene; Aspergillus fumigatus; infection;
 XX cancer; contamination; biofilm; antibody; immune response; ds.
 OS Aspergillus fumigatus.
 XX
 PN WO200286090-A2.
 PD 31-OCT-2002.
 XX
 PF 23-APR-2002; 2002WO-US013142.
 XX
 PR 23-APR-2001; 2001US-0285697P.
 XX 27-APR-2001; 2001US-0287066P.
 PR 05-JUN-2001; 2001US-0295890P.
 PR 09-JUL-2001; 2001US-0303899P.
 PR 31-NOV-2001; 2001US-0316362P.
 XX
 PA (ELIT-1) ELITRA PHARM INC.
 XX
 PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
 DR WPI; 2003-093124/08.
 XX
 PT New purified or isolated nucleic acids of essential genes of Aspergillus
 PR fumigatus, useful for treating or preventing infections by A. fumigatus,
 PT or for treating a non-infectious disease in a subject e.g. cancer.
 XX
 PS Disclosure; Page; 175pp; English.
 XX
 XX The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as A. fumigatus, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for
 CC expressing recombinant protein for characterization, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of A.
 CC fumigatus to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organisms invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence

CC factors. This polynucleotide sequence represents one of the essential
 CC genes of Aspergillus fumigatus of the invention
 XX
 SQ Sequence 1455 BP; 318 A; 406 C; 419 G; 312 T; 0 U; 0 Other;

Query Match 4.8%; Score 62.6; DB 8; Length 1455;
 Best Local Similarity 46.7%; Pred. No. 3.4e-06;
 Matches 375; Conservative 0; Mismatches 389; Indels 39; Gaps 4;

OY 387 AAAAATTTCTTACATTCGTCGTCATTCACCAAGTGAAGAAAGCCGATTAGTTC 446
 DB 513 AAGCATTTCTGATGACTTTGTTGCTTCGCTTCAGAGGACACAAAGAGCCTCGCTC 572
 OY 447 ATCGGACAGATTAGTCACTGTTGCCAAGTTTATATCCCA-----TTTATCCC 500
 DB 573 TTGAGCCGCTTGGTCACTGCGCTGATGTCGTCCTCTGCTATCCACCGTACTATGCAAC 632
 OY 501 CAATGTTATCAATGATGAATTAAGATATTTTGGACAGCTTGCACGATTTGACATTTGTA 560
 DB 633 CGACGATCTCGATCCAGGCGCGCAACGCTTCAATGAGCCGCGGCGCCACTGCGC 692
 OY 561 TGCCCAAAAAAGATAGATCTGGGTTGATGTTGCACTGCAATTATTAATGTTGATTTGT 620
 DB 693 TGCCGAGGTTAAAGTCGGCTCCGCTTGGATGTTGACACTGCTTACCGCTCCTGCT 752
 OY 621 ATATGAAAGATTTCAGCCAGCTTTGATTAATGACGTGTTCA---GGTTCTGAAAGTG 676
 DB 753 CTACAGACGCTTCTCCCTCGATTCGATTCGATGAGGTGACGCTGCTTCCAGGCTT 812
 OY 677 ATCTGAAAGTTCCCAACAGATGTAAGAAATTTGAAAGTAACTGGG-----A 728
 DB 813 CGAAGAGGGGTTGTTCCGATCTGTAAGAGAGCGCCACCCCAAGCTTCGTGGAGACCGA 872
 OY 729 ATTCAACATGAAGATGATATACATTACATAGCAATCAAGTATTAATGAGTGAAGTCA 788
 DB 873 ATGTTAATCTTCGCGCATGAAGCTCCCGCGGAATGAATGCTCTGTGAGCTGCA 932
 OY 789 GGGTGGCTCAGAAACCCCAATTTGATCAAGATCTCCATGAGAAAGGAAAGCC 848
 DB 933 ATGTGTTTCGACACTCATCCATGATGAGAAAGTTTGGAGTGGCGGAAACGAAATCA 992
 OY 849 AGAAGAAAGCTCTGTGTTGATGACCACTTAATATGTCACATTTACAGTTATGAAGA 908
 DB 993 GAAAGAACCCGATATGCTCTGGGGGCTCTGCAATCGAACACGAGACTTCGCTGGA 1052
 OY 909 ATTGAGGAAATGCGCTGAAAAATACAGCTCAGACCCAGAGACTTATATTAAGAGTTGA 968
 DB 1053 ACTCAGAGCGTTGGACAGAGCCGGAAGAACATCTCTCAGTGACTTTGAAAAATGTCG 1112
 OY 969 TCATTCTGTGAGCTTTGACTGTTGCGATTGAACATCGAAAAAGGTTTCAAGCAAT 1028
 DB 1113 CACTATATTTGAGCGCTCG-----CGTAAACACATCCGTTCCAT 1151
 OY 1029 AACACAAAATACAGGTTTCCAAATGAACCTGATGCCAAACCGATTGGACCGTTG 1088
 DB 1152 GACTCAAAAAGTCGATGTCATGACATGAGCGCGCTCCAAACGAGCTATTTGAGCTCT 1211
 OY 1089 TCAAGAGATCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1148
 DB 1212 GTCCAGAGCTGAGGGGTTCATCGATGATGATGATGATGATGATGATGATGATGATG 1271
 OY 1149 AGCTGATTTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1171
 DB 1272 TGTGCTCTCTCATCCAAAGCAATC 1294
 RESULT 35
 ID ABT20458 standard; DNA; 1509 BP.
 AC ABT20458;
 XX
 DT 16-APR-2003 (first entry)

XX XX Aspergillus fumigatus essential gene #2816.
 DE XX
 XX XX Fungicide; cytosstatic; essential gene; Aspergillus fumigatus; infection;
 KW cancer; contamination; biofilm; antibody; immune response; de.
 XX XX
 OS Aspergillus fumigatus.
 XX XX
 PN WO200286090-A2.
 XX XX
 PD 31-OCT-2002.
 XX XX
 PF 23-APR-2002; 2002WO-US013142.
 XX XX
 PR 23-APR-2001; 2001US-0285697P.
 XX XX
 PR 27-APR-2001; 2001US-0287066P.
 XX XX
 PR 05-JUN-2001; 2001US-0295890P.
 XX XX
 PR 09-JUL-2001; 2001US-0303899P.
 XX XX
 PR 31-AUG-2001; 2001US-0316362P.
 XX XX
 PA (ELIT-) ELITRA PHARM INC.
 XX XX
 PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM,
 XX XX
 DR WPI, 2003-093124/08.
 XX XX
 PT New purified or isolated nucleic acids of essential genes of Aspergillus
 PT fumigatus, useful for treating or preventing infections by A. fumigatus,
 PT or for treating a non-infectious disease in a subject e.g. cancer.
 XX XX
 PS Disclosure; Page: 175pp; English.
 XX XX
 CC The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as A. fumigatus, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for
 CC expressing recombinant protein for characterization, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of A.
 CC fumigatus to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This polynucleotide sequence represents one of the essential
 CC genes of Aspergillus fumigatus of the invention
 XX XX
 SQ Sequence 1509 BP; 338 A; 418 C; 429 G; 324 T; 0 U; 0 Other;
 XX XX
 Query Match 4.8%; Score 62.6; DB 8; Length 1509;
 Best Local Similarity 46.7%; Pred No. 3.5e-06;
 Matches 375; Conservative 0; Mismatches 389; Indels 39; Gaps 4;

QY 501 CAATGTTATCAATGCAATTAAGATATTTTGGACAACGTTGACAGATTGACATTTGTA 560
 DB 687 GCACGATCTCGGTCAGGCGCGGACAGCTTACATCTGGGCCAGGCGGCCACTGCGC 746
 QY 561 TGCCCAAAAAAAGATAGATCTGGTTGATTTGATTCGAACTGCAATTTATGCTGATTT 620
 DB 747 TGCCGAGGATAAAGTCGAGTCGAGCTCGAGTTGACACTGCAATTTACGGCTCTGCT 806
 QY 621 ATATGAGATTTTACGACGCTTTGATTAATGACGCTTTTCA-----GGTTAGAAAGTG 676
 DB 807 CTACGAGCGCTTCTCCCTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 866
 QY 677 ATCTGAGAGTTTCCCAAGAGTTGAAATTTGAAATTTGAAATTTGAAATTTGAAATTTG 728
 DB 867 GGAAGAGGAGTTGTTCCGAGATGTAGAGAGAGCGGACCCCGAGATTCGTTGGAGACCGA 926
 QY 729 ATTCAAAATGAAAGATGATCAATTTCAATCGGAATCAAGTTATTTAATGGTGAGCTCAA 788
 DB 927 ATGTCTAGACTTCCGAGATGAAAGCTCCCGCGAATGCAAAATGATCTCTGTGAGAGTCA 986
 QY 789 GGGTGGCTCAGAAACACCAATTTGATACAGTACTCCAAATGAGAAAGAAAGCC 848
 DB 987 ATGTGATTCGACGATCTCCATTCATGTAGAAAGTTTGGAGTGGCGGAAACAGATCA 1046
 QY 849 AGAAGAAAGCTCTGTGTGTATGATGACAGCTTAAATAGTCCAAATTTATAGAGA 908
 DB 1047 GAAAGAGCGATATGCTCTGGGCGCTGTGCAATGCAACAGAGACTTCGCTGGA 1106
 QY 909 ATTGAGGAAATGCGTGAATTAATGACTCAGACCCAGAGACTTATTTAAGAGATTAGA 968
 DB 1107 ACTCAGAGCGTTGGACAGAGCCCGAGCAATCTCTCAGTGAATTTGAAATGTCG 1166
 QY 969 TCATCTGTGAGCTTTGACGTTGCGATTAAAGATCAGAAAGGTTACAGACTT 1028
 DB 1167 CACTATATATGAGCCTG-----CTAACACATCGTTTCAAT 1205
 QY 1029 AACCAAAATGAGAGTTTCAATTTGAACCTGATGTCAAACCGATTGTTGACCGTTG 1088
 DB 1206 GACTCAAAAGTGGATGATGTCGATGAGCGCGCTCAACAGAGCTACTGAGCGCTCT 1265
 QY 1089 TCAAGAGATCTCTGTTGT 1148
 DB 1266 GTCCAGAGTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1325
 QY 1149 AGCTGATTTAGCTTGAATATC 1171
 DB 1326 TGTGCTCTCATCAAGCAATC 1348
 RESULT 36
 ID ABT18642 standard; DNA; 1509 BP.
 XX ABT18642;
 AC 16-APR-2003 (first entry)
 XX
 DT Aspergillus fumigatus essential gene #1000.
 DE
 XX Fungicide; cytosstatic; essential gene; Aspergillus fumigatus; infection;
 KW cancer; contamination; biofilm; antibody; immune response; de.
 XX
 OS Aspergillus fumigatus.
 XX
 PN WO200286090-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 23-APR-2002; 2002WO-US013142.
 XX
 PR 23-APR-2001; 2001US-0285697P.
 XX
 PR 27-APR-2001; 2001US-0287066P.
 XX
 PR 05-JUN-2001; 2001US-0295890P.

PR 09-JUN-2001; 2001US-0303899P.
 PR 31-AUG-2001; 2001US-0316362P.
 PA (ELIT-) ELITRA PHARM INC.
 PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
 DR WPI; 2003-093124/08.
 XX New purified or isolated nucleic acids of essential genes of *Aspergillus*
 PT fumigatus, useful for treating or preventing infections by *A. fumigatus*,
 PT or for treating a non-infectious disease in a subject e.g. cancer.
 XX Disclosure; Page; 175pp; English.
 PS
 XX The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or inhibit formation of an object
 CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
 CC expressing recombinant protein for characterisation, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of *A.*
 CC *fumigatus* to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This polynucleotide sequence represents one of the essential
 CC genes of *Aspergillus fumigatus* of the invention
 XX
 SQ Sequence 1509 BP; 338 A; 418 C; 429 G; 324 T; 0 U; 0 Other;
 Query Match 4.8%; Score 62.6; DB 8; Length 1509;
 Best Local Similarity 46.7%; Pred. No. 3.5e-06;
 Matches 375; Conservative 0; Mismatches 389; Indels 39; Gaps 4;
 QY 387 AAAAAATTCTTACCATCTCGTGCCATTACGAGTGAAGAAAGCGGATTAGGTTTC 446
 DB 567 AAGCAAGATTGCGGACTTTGGTGGCTTCAGAGAGGACACAGACAGGCGCTGGGCTC 626
 QY 447 ATGGGCGAGATTAGTGTGCTGTCGACAACTTTATATCCCA-----TTTATCCC 500
 DB 627 TTGAGCGCGCTTGGTCACTGCTGCTGCTCTCTGATCCACCTGATGCAAC 666
 QY 501 CAATGTATCAGTACGAAATGAATATTTGCAACAAGTTCAGATTCAGATTGTTA 560
 DB 687 CGAGATCTGGTCGAGCGGCGGACAGCTTCAACATCTGGCCAGGCGGCGCCACGCGC 746
 QY 561 TGCCCAAAAAAAGATAGATCTGGGTTGATGTTGCACTGCAATTTATGCTGATTTG 620
 DB 747 TGGCAGAGGTAAAGTCCGGGCTCGGCTTCGATGTTGACGCTTTCACGCTTCCTGCTC 806
 QY 621 ATATGAAGATTTCAGCAGCTTTGATTAATGACGTTCCTCA-----GTTCTGAAGAGTG 676
 DB 807 CTAGAGAGCTTCTCCCTCGATTCGAAATCAGTGGGTGACCTGCTGTTCTCCAGCTT 866
 QY 677 ATCCTGAAGATTCCCGACAGATTTGAATAAATTGATTTGAAGTAACTAGTGG-----A 728
 DB 867 CGAAGACCGGTGTTCCGGAATCGTAGAGAGCGCCGACCCAGCATCCGTGGGACACCGA 926
 QY 729 ATTCAAAATGAAGATGTACATTACCATACGGAATCAAGTTATTAATGGGTGACGTCAA 788

DB 927 ATGTCTAGACTTCGGCATGAGAGCTCCCGCGAATGCAAAATGCTCCTGTGATCGTCA 986
 QY 789 GGGTGGCTCAAAAACCCCAATTGGATACAGAGTATCTCCAAATGAAAAAAGAACCC 848
 DB 987 ATGTGTTCCGAGCTTCATTCATGTGAGAAAAGTTTGGAGTGGCGGAAACAGATCA 1046
 QY 849 AGAAGAAGCTCTGTTGTGTATGACAGCTTAATAGTCCAAATTTACGTTTATGAAGA 908
 DB 1047 GAAGGAAGCGATATGCTCTGGGCGCTGTGCAATCGAACAGAGAGAGCTTGCCTGGA 1106
 QY 909 ATTGAGGAAATGCGTGAATAAATACGCTGACAGCCGAGACTTATTTAAAGTTAGA 968
 DB 1107 ACTCAGACGCTTGGCACAGAGCCCGAGCAACTCTCACTTCACTTAAATTTCCG 1166
 QY 969 TCATTCTGTGAGCTTGTGACTGTGGATTAAGAACATCGAAGAAAGGTTACAGCAAT 1028
 DB 1167 CACCTATATTCAGGCTCG-----CGTAAACATCCGTTCCAT 1205
 QY 1029 AACCAAAAATCAGAGGTTCCAAATTGAACCTGATGCCAAACCAATTGTTGACGTTG 1088
 DB 1206 GACTCAAAAGTCGATGTCCCAATCGAGCGCGCTCAAAACGAGCTACTTACGCTCT 1265
 QY 1089 TCAAGAGATTCTGTTGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1148
 DB 1266 GTCCGAGCTGAGGAGGTTCATCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1325
 QY 1149 AGCTGATTATGTTGGAAATC 1171
 DB 1326 TGTGCTCTCATCAAGCAATC 1348
 RESULT 37
 ABT18048
 ID ABT18048 standard; DNA; 3508 BP.
 XX
 AC ABT18048;
 XX
 XX 16-APR-2003 (first entry)
 DT
 XX
 DE *Aspergillus fumigatus* essential gene #406.
 XX
 XX Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;
 KW cancer; contamination; biofilm; antibody; immune response; ds.
 XX
 XX *Aspergillus fumigatus*.
 OS
 OS WO200286090-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 23-APR-2002; 2002WO-US013142.
 XX
 PR 23-APR-2001; 2001US-0285697P.
 PR 27-APR-2001; 2001US-0287066P.
 PR 05-JUN-2001; 2001US-0295890P.
 PR 09-JUL-2001; 2001US-0303899P.
 PR 31-AUG-2001; 2001US-0316362P.
 XX
 XX PA
 XX (ELIT-) ELITRA PHARM INC.
 XX
 PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
 DR WPI; 2003-093124/08.
 XX
 XX New purified or isolated nucleic acids of essential genes of *Aspergillus*
 PT fumigatus, useful for treating or preventing infections by *A. fumigatus*,
 PT or for treating a non-infectious disease in a subject e.g. cancer.
 XX
 PS Disclosure; Page; 175pp; English.
 CC The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of

CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
 CC expressing recombinant protein for characterization, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of *A.*
 CC *fumigatus* to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This polynucleotide sequence represents one of the essential
 CC genes of *Aspergillus fumigatus* of the invention

XX
 SQ Sequence 3508 BP; 857 A; 916 C; 883 G; 852 T; 0 U; 0 Other;

Query Match 4.8%; Score 62.6; DB 8; Length 3508;

Beet Local Similarity 46.7%; Pred. No. 4.9e-06;

Matches 375; Conservative 0; Mismatches 389; Indels 39; Gaps 4;

QY 387 AAAAAATTCTTACATTCGTCGTCATACCGAAGAGGAAACCGGATTAGTTC 446
 DB 1566 AAGCATTCGCGAGATTGTTGTTGCTTACAGAGGACACAAACAGCGCTGGCTC 1625
 QY 447 ATCGGAGATTAGTCTCAGTTGTTGCCAAGTTATTATCCCA-----TTTATCCC 500
 DB 1626 TTCACGCGGCTTGTCACCTGCTGTCGTCCTGTCATCACCGTACTATGCAAC 1685
 QY 501 CAATGTTATCAGTACGAAATTAAGATTTTTCGACACGTTGCAACATTTGACATTTGTA 560
 DB 1686 CGACGATTCGCTGCGAGCGCGGACAAAGCTTCAATTCGCGCGCGCCCACTGCGC 1745
 QY 561 TSCCCAAAAAAGATAGATCTGGGTTTATGTTGCACTGCAATTTATGTCGTTATG 620
 DB 1746 TCCGCAAGGTTAAGTCGGGCTCCGGCTTCATGTTGACGCTGCATTTACGCTCTCT 1805
 QY 621 ATATGAAGATTTCAGCCAGCTTTGATTAATGACGTTTCA-----GGTTCTAGAAAGTG 676
 DB 1806 CTACAGACGCTTCTCCCTCGATTCGATCAGTGGGAGAGCGTGTTCAGGCTT 1865
 QY 677 ATCCGAGAGATTCCCCACAGAGTTAAAAAATTGATTGAAGTAAGTGGG-----A 728
 DB 1866 CGAAGAGCGGTTGTTCCGATGTAAGAGACGCCGACCCCGACATCCCTGGGACACCA 1925
 QY 729 ATTCAAAACATGAAGATGATACATTACATACGGAATCAAGTTATTATGGTGAAGTCA 788
 DB 1926 ATGTCTAGACTTCGGCATAGACTCCCGCGGAATGCAAAATGCTCTCTGACGCTTCA 1985
 QY 789 GGGTGGCTCAGAAACACCCAAATGGTATACAGAGTACTCCATGAGAAAAAGGAAAGCC 848
 DB 1986 ATGTGTTTCGACAGATTCATCATGATGTAAGAAAGTTTGGATGGCGGAAACAGATCA 2045
 QY 849 AAAAAAGAGCTGTGTTGTATGACACAGTTAATAGTGCATTTACATTTATGAAGA 908
 DB 2046 GAAGAGGCGATATGCTCTGGGCGCTTGCATGCAACAGAGAGCTTGCCTTGA 2105
 QY 909 ATTGAGGGAATGCGGTAATAATATGAGTCAAGCCAGAGACTTATATTAAGAGTTAGA 968
 DB 2106 ACTCAGAGCTTGGCAGACAGCCGGAACAACTATCTCAGTGAAGCTTTGAATATGTCG 2165
 QY 969 TCATTTCTGTGAGCCTTTGACTGTGGCATTAAGAACATCAGAAAAAGGTTACMACATT 1028

DB 2166 CACTATATTCAGCGCTG-----CGTACACATCCGTTCCAT 2204
 QY 1029 AACACAAAATCAGAGTTCATTAACCTGATGTCGAACCCAGTTGTGACGGTTG 1088
 DB 2205 GACTCAAAAGTCGAGATGTCCTCAATCAGCGCGCTCAACCGACTTACGCTCT 2264
 QY 1089 TCAAGAGATTCCTGTTGTTGGTGTGTTGTTCCAGGTGCTGTGATACATCAAT 1148
 DB 2265 GTCCGAGCTGAGAGGATGTCATGCTGTGTTGTTCCAGAGACAGGGGCTAGCAGCCAT 2324
 QY 1149 AGCTGATATGATGTTGAAAAATC 1171
 DB 2325 TGTGCTCTCATCAACAGCAATC 2347

RESULT 38

ID ABT19862 standard; DNA; 3509 BP.

AC ABT19862;

DT 16-APR-2003 (first entry)

DE *Aspergillus fumigatus* essential gene #2220.

KM Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;
 KW cancer; contamination; biofilm; antibody; immune response; ds.

XX *Aspergillus fumigatus*.

PN WO200286090-A2.

PD 31-OCT-2002.

PF 23-APR-2002; 2002WO-US013142.

PR 23-APR-2001; 2001US-0285697P.

PR 27-APR-2001; 2001US-0287066P.

PR 05-JUN-2001; 2001US-0295890P.

PR 09-JUL-2001; 2001US-0303899P.

PR 31-AUG-2001; 2001US-0316362P.

XX (ELIT-) ELITRA PHARM INC.

PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

DR WPI; 2003-093124/08.

XX New purified or isolated nucleic acids of essential genes of *Aspergillus*

PT *fumigatus*, useful for treating or preventing infections by *A. fumigatus*,

PT or for treating a non-infectious disease in a subject e.g. cancer.

XX Disclosure; Page; 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
 CC expressing recombinant protein for characterization, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of *A.*
 CC *fumigatus* to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding

interaction. The polypeptides may be used to raise antibodies or to elicit immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case of virulence factors. This polynucleotide sequence represents one of the essential genes of *Aspergillus fumigatus* of the invention

XX Sequence 3509 BP, 857 A, 917 C, 883 G, 852 T, 0 U, 0 Other;

Query Match 4.8%; Score 62.6; DB 8; Length 3509;

Best Local Similarity 46.7%; Pred. No. 4.9e-06; Matches 375; Conservative 0; Mismatches 389; Indels 39; Gaps 4;

```

QY 387 AAAAAATTCTTTACCATTCGTCGCCATTACCGAAGTGGAAAAAGCCGATTAGTTTC 446
DB 1567 AAGCAGATTCTGTGACCTTGTGTTCCCTTCAGAGAGGACACAAGACAGGCTGGGCTC 1626
QY 447 ATCGGAGAGATTAGTGTGAGTTGGCCACAGATTATATCCCA-----TTTATCCC 500
DB 1627 TTGAGCGGCTGTGCTACTGCCCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1686
QY 501 CATGTTATCAGTACGATTAAGATATTTTGGACAACGTTGACAGATTGACATTGTTTA 560
DB 1687 CGACGATCTCGGTCCAGCGCCGACAAAGCTTCAATCTGCCCCAGGCGGCCACTGCGC 1746
QY 561 TGGCCCAAAAAAGATAGATCTGGGTTTATGTTGAACTGCAATTATGTCATTGTT 620
DB 1747 TGGCAGGGTAAAGTCGGGTCCGGCTTCATGTGGAGTGCCTATTTAAGGCTCTCTGT 1806
QY 621 ATATAGAGATTTTACGACGCTTTGATTAATGACGTTTCA-----GTTCTAGAAAGTG 676
DB 1807 CTACAGACGCTTCTCCCTCCATTCGAAATCAGTGGGTGACGCTGTTCTCCAGGCTT 1866
QY 677 ATCTGAGAAGTTCCTCCACAGAGTTGAAAAAATTGATGAAGTAATCTGG-----A 728
DB 1867 CGAAGAGCGGTGTGTTCCGATCTAGAGACGCGCACCCCGACATCCGTGGGACACCGA 1926
QY 729 ATTCAAAACATGAAGATGATACATTACCAACGGAATCAGATTATTAATGGGTAGCTCA 788
DB 1927 ATGTCTAGACTTCGGATGAAGCTCCCGCGAATGCAAAATGCTCTGTGACGTCGA 1986
QY 789 GGGTGGCTCAGAAACACCCAAATTTGATNACAGATCTCAATGAGAAAAAGGAAAAAGCC 848
DB 1987 ATGTGTTGGCAATCTCCATCCATGATGAGAAGATTTTGGAGTGGGAGAAACAGATCA 2046
QY 849 AGAAGAAAGCTCTGTGTGTATGACCACTTAATAGTCCCAATTTCAGTTTATGAAGA 908
DB 2047 GAAGGAAGCCGATATGCTCTGGGGCGCTCTGCAATCGAACACAGAGACTTGCCTGGA 2106
QY 909 ATTGAGGGAATATCGTGAAATAATACGACTCAGACCCAGACTTATATTAAGATTAGA 968
DB 2107 ACTCAGACGCTTGACAGAGCCCGGAGAACATCTCTCAGTGACTTTGAAAAATGTCG 2166
QY 969 TCATTCTGTGAGCCTTTGACTGTGTGCGATTAAAGAACATCAGAAAAAGGTTTACAAG 1028
DB 2167 CACCTATATCTCAAGGCTCG-----GTAACCAATCCGTTCCAT 2205
QY 1029 AACACAAAATCAGAGTTTCAATTGAACCTGATGTCAAACCAAGTTTGGACCGTTG 1088
DB 2206 GACTCAAAAGTCGATGTCCTCAATCGACGCGCGCTCAAAACCGAGTTACTTGAAGCTCT 2265
QY 1089 TCAAGAGATTCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1148
DB 2266 GTCCGAGCTGAGGGGTCTACTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2325
QY 1149 AGCTGATTAGTGTGAAAAATC 1171
DB 2326 TGTGCTCTCTATCCAGACATC 2348

```

RESULT 39
ABZ51452

```

ID ABZ51452 standard; cDNA; 869 BP.
AC ABZ51452;
XX
DT 28-MAR-2003 (first entry)
XX
DE Aspergillus oryzae polynucleotide SEQ ID NO 565.
XX
KM Aspergillus oryzae; fermentation; fungus; industrial; EST;
KM expressed sequence tag; gene; ss.
XX
OS Aspergillus oryzae.
XX
PN MO200279476-A1.
XX
PD 10-OCT-2002.
XX
PF 22-MAR-2002; 2002MO-IB000890.
XX
PR 30-MAR-2001; 2001JP-00098371.
XX
PA (NAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (NARE-) NAT RES INST BREWING.
PA (NORO) NAT FOOD RES INST MIN AGRIC.
XX
PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
DR WPI, 2003-046817/04.
XX
PT Detection of expression of specific Aspergillus genes for monitoring the
PT fermentation and growth conditions of the fungus, using DNA probes.
XX
PS Claim 1; SEQ ID NO 565; 48bp + Sequence Listing; Japanese.
XX
CC The invention relates to a polynucleotide having any of 6006 specific
CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridizing
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 869 BP, 193 A, 226 C, 239 G, 211 T, 0 U, 0 Other;

```

Query Match 4.1%; Score 53.2; DB 8; Length 869;

Best Local Similarity 53.5%; Pred. No. 0.00092; Matches 137; Conservative 0; Mismatches 113; Indels 6; Gaps 1;

```

QY 390 AACATTCTTTACCATTCGTCGCCATTACCGAAGTGGAAAAAGCCGATTAGTTTATC 449
DB 339 AAGATTCGGAACCTTCGGGTTCCTCTTCAAGAGCCCAAGACAGGCTAGGTCTTC 398
QY 450 GCGAGATTAGTGTGAGTTGTTGCCAAGTTTATATCCATTATC-----CCCA 503
DB 399 TGGGGCTCTAGTACTGCCCTTAGTATCACTCTGTTATCAACCGTACCTGACGCTGA 458
QY 504 TGTATCAGTACGATTAAGATATTTTGGACAACGTTGACAGATTGACATTGTTATGC 563
DB 459 CGACCTTGAGGCTTCTCGTGAACGTTCAATACTTGACAGGCTGCCACATGTGCTGC 518
QY 564 CCAAAAAAGATAGATCTGGGTTTGAATTTGCAACTGCAATTATGATGCTGATATGATA 623
DB 519 TCAAGGTAAAGTGGGATCCGGGTTGATGTGCTGCTATCTACGAGCTCTTGCTATA 578
QY 624 TAGAAGATTTCAGCCA 639
DB 579 TGGCGATTCTCCCA 594

```

```
RESULT 40
ADA71938/C
ID ADA71938 standard; DNA; 2000 BP.
XX
XX ADA71938;
AC ADA71938;
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX Rice gene, SEQ ID 5263.
DE
XX
XX Plant, bacterial infection, fungal infection, viral infection, rice;
KM gene; ds.
XX
XX Oryza sativa.
OS
XX WO2003000898-A1.
PN
XX 03-JAN-2003.
PD
XX 22-JUN-2001; 2001WO-1B001105.
PF
XX 22-JUN-2001; 2001WO-1B001105.
PR
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
DR
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5263; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
SQ
Query Match 3.7%; Score 48.2; DB 8; Length 2000;
Best Local Similarity 6.8%; Pred. No. 0.028;
Matches 52; Conservative 372; Mismatches 340; Indels 3; Gaps 2;
QY 18 TGCACCTGGAAAGACATTTCTGCTGGTGGATATTTGTTCTTGAGCAATTATGATGAC 77
DB 1018 TGMATTTTSSMWTYATMMKKTKTMTATYSTWKMTAYKRAVWRSRKTWCTGGKRM 959
QY 78 TTATGTGACAGCATTTGTCATCAAGATGCATGATGATTAACCAAAAGAACCATTT 137
DB 958 ATYCGTKMMAAGRWBRBAMWCYCCMKMKWKTSCMMWKYWTWSCWYTMWMMGAKRYAY 899
QY 138 GAAAGATCTGAATCAAAATTTCTTACCCCAATTTGCAACGGAATGGAATATCA 197
DB 898 AWRRRRWYTKWSWRMYWTMTWTMTATMTATMTATMTATMTATMTATMTATMTATMT 839
QY 198 CATATCATCAATATGACAGAGCCGAGAGATTCAGTACGATCAATCAATTTTGA 257
DB 838 CAKCYTKMAATKMTTWTACAMRATSRWRBAMGMRKRYKMKRAYWRRWRCKAGMARW 779
QY 258 GGCAACTATATTCATCGTTTGTATATTCATCAACCGACGAGCATTTGATCTTGAAT 317
```

```
DB 778 MRS-RYRMKMKKYATATRYWKMAMTMMWSWRMKSRYRMWSGMRMWSAMRYCSRKCKAK 720
QY 318 CATCATTTTACTGACCCCTGATATCATTCACAGAAGATACTGAAACAAACATCTCTC 377
DB 719 TKYASARWTKAASYSRYRRRWYKWKGTTRYRYWRSCRTTRAMSGRRKAGASMS 660
QY 378 GAAT--GGAGAAAAAATCTTTTCAATTCCTGTCGTCATTACCGAATGGAAGAACACC 435
DB 659 CMWYRGAASWYKSKYSCAKCKTKRWYVTSYSTGYMYGMYSYKSMWTSKMSYGM 600
QY 436 GGATTAAGTTCATCGGACGATTAAGTTCAGTTGTCAGTTCGCCAAGATTATATCCATTT 495
DB 599 TCTMYTSMKGSYTRSRKMGWSGMSRWYRMWKKRKYRMYKMKCTWRRCMYRWGYT 540
QY 496 ATCCCAATGTTATCAGTACGATGAATTAAGATATTTTGACACAGTTCACAGATTGCACAT 555
DB 539 MTTTSRSMYTKRYKARTSKRYRMWYKRYKCYIYGYMKKSYMMRYGICAKCKK 480
QY 556 TGTATGCCCCAAAAAAGATAGATCTGGTTTGATGTTGCAACTGCAATTTATGCTCTG 615
DB 479 CYAMCMKAAYSQMMWYTRKYSKMMRMSYKMMWSMWYKCKRSMKYGAKGCGYCKMTTYS 420
QY 616 ATTTATATAGAAGATTTTCAAGCCTTTGATTAATGACGTTTCAGTCTTACAGT 675
DB 419 YGYMKWYTYMGSYKYSRCYKRYRMWYKGMWYMYSAYSMMTWYVYAKYKWKYKRR 360
QY 676 GATCCGGAAGATTCGCCACAGAGTTGAAAAAATGATTGAAGAAAGTGGGAATTCAA 735
DB 359 GTMSWYGSYKRYKCTTMCYMKCMRCYRWRKMKKTKYSKRCYCRVATCYCCCTYKR 300
QY 736 CATGAAGATGTACATTAACATACGAAATCAAGTTATTTATGAGTGA 782
DB 299 GMYSRSMWRTAGKMKMSWSRHCYSYSWYKMKMKSYMSYSGW 253
```

Search completed: January 26, 2005, 04:43:00
Job time : 496 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: January 26, 2005, 06:40:30 ; Search time 3871 Seconds
(without alignments)
15869.109 Million cell updates/sec

Title: US-10-069-062-6
Perfect score: 1299
Sequence: 1 atgtcaaaagcatcttgctc.....aagactataggtttataa 1299

Scoring table: OLIGO_NUC
Gapop_60.0, Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_rsb: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1299	100.0	1299	6	AX087879 Sequence
2	1299	100.0	1763	6	AX087878 Sequence
3	891	68.6	1299	6	AX240804 Sequence
4	891	68.6	1299	6	AX488723 Sequence
5	421	32.4	547	6	AX087874 Sequence
6	334	25.7	577	6	AX087876 Sequence
7	65	5.0	65	6	AX485723 Sequence
8	65	5.0	25	6	AX241110 Sequence
9	25	1.9	25	6	AX087877 Sequence
10	24	1.8	36	6	AX087884 Sequence
11	23	1.8	110000	10	AE014174_2
12	23	1.8	201532	2	AC115701 Mus muscu
13	22	1.7	22	6	AX241232 Sequence
14	22	1.7	22	6	AX487723 Sequence
15	22	1.7	58879	2	AC090350 Homo sapi
16	22	1.7	153598	2	AC149142 Xenopus t
17	22	1.7	171133	9	AC103846 Homo sapi
18	22	1.7	185418	2	AC101231 Homo sapi
19	22	1.7	188270	9	AC104393 Homo sapi

C	20	1.7	192347	9	AC007614
C	21	1.6	21	6	AX087875
C	22	1.6	313	6	CO661526 Sequence
C	23	1.6	479	6	AX396511 Sequence
C	24	1.6	530	6	AX397039 Sequence
C	25	1.6	771	6	BD020934
C	26	1.6	771	6	BD100872
C	27	1.6	1444	6	CQ171294
C	28	1.6	1596	4	AF068483 Oryctolag
C	29	1.6	1685	4	HUMHTR3A
C	30	1.6	2010	9	HUMTCR20
C	31	1.6	2562	9	AF385084
C	32	1.6	2647	9	AB063318 Homo sapi
C	33	1.6	2837	6	CQ492936 Sequence
C	34	1.6	23130	2	AC020045
C	35	1.6	38875	9	AC092101 Homo sapi
C	36	1.6	115815	8	AC114896
C	37	1.6	128524	2	AC012253
C	38	1.6	128685	10	AL772361
C	39	1.6	134105	5	AC096845
C	40	1.6	145042	10	AL732425
C	41	1.6	152927	9	AC004895
C	42	1.6	153174	2	AC135904
C	43	1.6	154791	2	AC069393
C	44	1.6	164669	2	AC140966
C	45	1.6	174394	2	AC112725 Drosophi
C	46	1.6	183338	2	AC083775
C	47	1.6	187772	3	AC104628
C	48	1.6	188914	9	AC093744
C	49	1.6	189976	3	AC023710
C	50	1.6	191291	2	AC148313
C	51	1.6	197639	9	AC009554
C	52	1.6	198684	2	AC148309
C	53	1.6	201932	5	BX897729
C	54	1.6	203816	2	CR407544
C	55	1.6	215435	10	AL662838
C	56	1.6	215972	2	AC146017
C	57	1.6	217022	2	AC146146
C	58	1.6	232385	2	AC133770
C	59	1.6	238009	2	AC103650
C	60	1.6	242109	2	AC128430
C	61	1.6	242467	2	BX927324
C	62	1.6	305376	3	AE003437
C	63	1.5	33	6	AX087883
C	64	1.5	630	6	AX505906
C	65	1.5	630	6	BR006211
C	66	1.5	919	8	AK117303
C	67	1.5	1476	8	AY113176
C	68	1.5	1535	3	SAR31MSA
C	69	1.5	1978	8	AY054231
C	70	1.5	2490	8	BT003013
C	71	1.5	2856	8	AY125524
C	72	1.5	3527	5	XU126256
C	73	1.5	3585	5	BC059339
C	74	1.5	14745	9	PTEN3
C	75	1.5	39487	9	AC148869
C	76	1.5	52717	8	AB019227
C	77	1.5	62580	8	AL355143
C	78	1.5	63710	2	AC084183
C	79	1.5	74560	9	HS433B8
C	80	1.5	89582	9	AC022429
C	81	1.5	100000	9	AP000093
C	82	1.5	100000	9	AP000197
C	83	1.5	107515	2	AC147403
C	84	1.5	107603	2	AC016661
C	85	1.5	110000	2	AC123648_2
C	86	1.5	110000	2	AC123879_1
C	87	1.5	110000	2	AC140127_2
C	88	1.5	113126	8	AC007576
C	89	1.5	115348	8	AL357042
C	90	1.5	116166	9	AC024109
C	91	1.5	118439	9	AL135905
C	92	1.5	120045	9	AL353052

AC007614	Homo sapi
AX087875	Sequence
CO661526	Sequence
AX396511	Sequence
AX397039	Sequence
BD020934	Novel gen
BD100872	Novel gen
CQ171294	Sequence
AF068483	Oryctolag
M94083	Human chape
L27706	Human chape
AF385084	Homo sapi
AB063318	Homo sapi
CQ492936	Sequence
AC020045	Drosophi
AC092101	Homo sapi
AC114896	Oryza sat
AC012253	Homo sapi
AL772361	Mouse DNA
AC096845	Takifugu
AL732425	Mouse DNA
AC004895	Homo sapi
AC135904	Rattus no
AC069393	Homo sapi
AC140966	Sus scrofa
AC112725	Drosophi
AC083775	Homo sapi
AC104628	Drosophi
AC093744	Homo sapi
AC023710	Drosophi
AC148313	Canis trogl
AC009554	Homo sapi
AC148309	Canis trogl
BX897729	Zebrafish
CR407544	Danio rer
AL662838	Mouse DNA
AC146017	Pan trogl
AC146146	Pan trogl
AC133770	Rattus no
AC103650	Mus muscu
AC128430	Rattus no
BX927324	Danio rer
AE003437	Drosophi
AX087883	Sequence
AX505906	Sequence
BR006211	Arabidops
AK117303	Arabidops
AY113176	Arabidops
M76496	Sarcocystis
AY054231	Arabidops
BT003013	Arabidops
AY125524	Arabidops
XU126256	Xenopus lea
BC059339	Xenopus lea
PTEN3	Human DNA
AC148869	Pan trogl
AB019227	Arabidops
AL355143	Human DNA
AC084183	Homo sapi
HS433B8	Human DNA
AC022429	Homo sapi
AP000093	Homo sapi
AP000197	Homo sapi
AC147403	Dasyatis n
AC016661	Arabidops
AC123648_2	Continuation (3 of
AC123879_1	Continuation (3 of
AC007576	Arabidops
AL357042	Human DNA
AC024109	Homo sapi
AL135905	Human DNA
AL353052	Human DNA

93	20	1.5 120112	9	AC124850	166	20	1.5 211843	2	AC145094	AC145094 Homo sapi
94	20	1.5 120527	8	AC141114	167	20	1.5 211864	2	AC139790	AC139790 Homo sapi
95	20	1.5 125529	2	AC121242	168	20	1.5 213388	2	AC138920	AC138920 Homo sapi
96	20	1.5 128039	2	AC010421	169	20	1.5 213867	2	AC141595	AC141595 Homo sapi
97	20	1.5 130630	2	AC010637	170	20	1.5 215130	5	AL954320	AL954320 Zebrafish
98	20	1.5 132403	2	AC148792	171	20	1.5 216082	10	AC107369	AC107369 Mus muscu
99	20	1.5 133561	9	AL951044	172	20	1.5 218336	9	AF067844	AF067844 Homo sapi
100	20	1.5 134091	2	AC023349	173	20	1.5 222040	2	AC098532	AC098532 Rattus no
101	20	1.5 134722	8	AC137923	174	20	1.5 226382	2	AC097295	AC097295 Rattus no
102	20	1.5 139121	2	AC135916	175	20	1.5 227089	2	AC079535	AC079535 Mus muscu
103	20	1.5 140971	2	AC145253	176	20	1.5 228353	2	AL445283	AL445283 Homo sapi
104	20	1.5 146032	2	AC145253	177	20	1.5 229447	2	AC123228	AC123228 Rattus no
105	20	1.5 152445	9	AC138410	178	20	1.5 233137	2	AC094489	AC094489 Rattus no
106	20	1.5 153265	9	AP000237	179	20	1.5 235468	10	AC105586	AC105586 Rattus no
107	20	1.5 153657	9	AC087145	180	20	1.5 239901	2	AC125552	AC125552 Rattus no
108	20	1.5 154311	9	AF107258	181	20	1.5 241629	2	AC136278	AC136278 Rattus no
109	20	1.5 154891	9	AC104131	182	20	1.5 242931	2	AC107478	AC107478 Rattus no
110	20	1.5 160582	9	AC105429	183	20	1.5 243366	2	AC150537	AC150537 Rattus no
111	20	1.5 160653	2	AC026306	184	20	1.5 243920	2	AC117886	AC117886 Rattus no
112	20	1.5 161460	9	AC093243	185	20	1.5 249520	2	AC096833	AC096833 Rattus no
113	20	1.5 162964	9	AL157879	186	20	1.5 252771	2	AC129788	AC129788 Rattus no
114	20	1.5 163031	9	AC011939	187	20	1.5 252311	2	AC135471	AC135471 Rattus no
115	20	1.5 163311	9	AC022016	188	20	1.5 262234	2	AC096148	AC096148 Rattus no
116	20	1.5 163911	2	AC138047	189	20	1.5 302422	1	AE017011	AE017011 Bacillus
117	20	1.5 164308	2	AC147402	190	20	1.5 31141	2	AC145103	AC145103 Homo sapi
118	20	1.5 164598	2	AC140143	191	20	1.5 340000	9	AP001706	AP001706 Homo sapi
119	20	1.5 165477	2	AC138938	192	19	1.5 206	9	HS75E1R	HS75E1R
120	20	1.5 165477	2	AC138938	193	19	1.5 247	11	BV088038	BV088038 sc1352_p3
121	20	1.5 167798	2	AC138947	194	19	1.5 255	11	BV088042	BV088042 sc1352_p3
122	20	1.5 168173	2	AC138947	195	19	1.5 399	11	BV116618	BV116618 PZA00783
123	20	1.5 168547	2	AC138947	196	19	1.5 406	11	BV116623	BV116623 PZA00783
124	20	1.5 169082	2	AC138933	197	19	1.5 414	11	BV116625	BV116625 PZA00783
125	20	1.5 170044	2	AC137696	198	19	1.5 452	3	AY363002	AY363002 Colieter
126	20	1.5 170264	2	AC137640	199	19	1.5 623	11	BV146882	BV146882 PZA02654-
127	20	1.5 170816	2	AC138411	200	19	1.5 646	11	BV146887	BV146887 PZA02654-
128	20	1.5 171705	2	AC139458	201	19	1.5 647	11	BV146880	BV146880 PZA02654-
129	20	1.5 175447	9	AC091607	202	19	1.5 648	11	BV146879	BV146879 PZA02654-
130	20	1.5 175942	2	AC140014	203	19	1.5 648	11	BV146888	BV146888 PZA02654-
131	20	1.5 176010	2	AC062274	204	19	1.5 656	11	BV146884	BV146884 PZA02654-
132	20	1.5 176189	2	AC133283	205	19	1.5 657	11	BV146881	BV146881 PZA02654-
133	20	1.5 177462	2	AC140012	206	19	1.5 657	11	BV146883	BV146883 PZA02654-
134	20	1.5 178464	2	AC140012	207	19	1.5 659	11	BV146886	BV146886 PZA02654-
135	20	1.5 178805	2	BX132260	208	19	1.5 661	11	BV146878	BV146878 PZA02654-
136	20	1.5 178932	2	AC138858	209	19	1.5 669	11	BV146885	BV146885 PZA02654-
137	20	1.5 182173	2	AC079563	210	19	1.5 669	11	BV146885	BV146885 PZA02654-
138	20	1.5 182991	2	AC121349	211	19	1.5 1107	8	AK067291	AK067291 Oryza sat
139	20	1.5 183155	9	BS000175	212	19	1.5 1242	8	MCU80070	MCU80070
140	20	1.5 183694	9	AC005412	213	19	1.5 1282	6	CQ492059	CQ492059 Sequence
141	20	1.5 183815	9	AL451084	214	19	1.5 1282	6	CQ497971	CQ497971 Sequence
142	20	1.5 184624	2	AC138868	215	19	1.5 1471	3	AY363004	AY363004 Colieter
143	20	1.5 185096	9	AL355140	216	19	1.5 2069	5	AE031243	AE031243 Gallus jap
144	20	1.5 185220	2	AC140173	217	19	1.5 2085	5	CH389521	CH389521 Oryza sat
145	20	1.5 186595	10	AC140053	218	19	1.5 2649	8	AK100281	AK100281 Oryza sat
146	20	1.5 187740	9	AC138409	219	19	1.5 3471	8	SCYGL049C	SCYGL049C Lactococc
147	20	1.5 187939	5	AL928672	220	19	1.5 4265	14	AF458777	AF458777 Lactococc
148	20	1.5 188179	2	AC139286	221	19	1.5 4532	14	AY029520	AY029520 FliJ dise
149	20	1.5 188352	2	AC139496	222	19	1.5 5249	10	AE006257	AE006257 Lactococc
150	20	1.5 188642	9	HS55C20	223	19	1.5 5774	10	MMU2188129	MMU2188129 S.pombe c
151	20	1.5 190194	9	AC068025	224	19	1.5 6490	8	SPBC1215	SPBC1215 Sequence
152	20	1.5 191957	9	AC015845	225	19	1.5 6733	6	CQ612740	CQ612740 Sequence
153	20	1.5 193309	2	AC138958	226	19	1.5 10769	6	CQ830738	CQ830738 Sequence
154	20	1.5 196263	2	AC138820	227	19	1.5 11137	6	AE003868	AE003868 Xylella f
155	20	1.5 196264	2	AC139780	228	19	1.5 12181	6	CQ612728	CQ612728 Sequence
156	20	1.5 199001	2	AL157880	229	19	1.5 13350	3	CHRG46N07	CHRG46N07 Caenorhab
157	20	1.5 201462	2	AC139283	230	19	1.5 14674	9	AL513490	AL513490 Human DNA
158	20	1.5 201536	9	AC063965	231	19	1.5 16336	9	AC034208	AC034208 Homo sapi
159	20	1.5 202679	2	AC141260	232	19	1.5 20370	8	AP006078	AP006078 Lotus cor
160	20	1.5 203221	10	AC102689	233	19	1.5 60464	9	AL359673	AL359673 Homo sapi
161	20	1.5 204533	2	AC116994	234	19	1.5 62871	2	AC124264	AC124264 Homo sapi
162	20	1.5 204617	2	BX324160	235	19	1.5 62871	2	AC124264	AC124264 Homo sapi
163	20	1.5 207430	10	AL772236	236	19	1.5 64036	2	AC040921	AC040921 Homo sapi
164	20	1.5 207459	9	AC137923	237	19	1.5 70574	9	AC017050	AC017050 Homo sapi
165	20	1.5 210040	2	AC107603	238	19	1.5 70574	9	AC017050	AC017050 Homo sapi

969	18	1.4	140955	2	AP001971
970	18	1.4	141017	8	AC037197
971	18	1.4	141175	2	AC012110
972	18	1.4	141293	2	AP004337
973	18	1.4	141380	2	AC134053
974	18	1.4	141380	2	AC134053
975	18	1.4	141380	2	AC134053
976	18	1.4	141659	10	AL805927
977	18	1.4	141959	2	AC087130
978	18	1.4	142001	10	BX294194
979	18	1.4	142094	9	HS154319
980	18	1.4	142227	2	AL162492
981	18	1.4	142512	4	AP005126
982	18	1.4	142659	8	AC089994
983	18	1.4	142687	8	AC079357
984	18	1.4	142694	2	AC069228
985	18	1.4	142813	2	AC090624
986	18	1.4	142952	9	AL391495
987	18	1.4	142974	2	AC128075
988	18	1.4	143018	8	AP003507
989	18	1.4	143130	8	AC104272
990	18	1.4	143223	9	AL162725
991	18	1.4	143335	9	AC025188
992	18	1.4	143851	9	AC006029
993	18	1.4	143864	2	AC146463
994	18	1.4	143965	8	AP002839
995	18	1.4	144074	8	AP003302
996	18	1.4	144102	8	AP004854
997	18	1.4	144177	9	AL513487
998	18	1.4	144213	2	AC132002
999	18	1.4	144252	8	AP005399
1000	18	1.4	144362	2	AC016120

ALIGNMENTS

RESULT 1
 AX087879
 LOCUS
 DEFINITION Sequence 6 from Patent WO0114533.
 ACCESSION AX087879
 VERSION AX087879.1 GI:13396872
 KEYWORDS
 SOURCE
 ORGANISM
 Candida albicans
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Mitosporic Saccharomycetales; Candida.
 REFERENCE
 1 Rosamond, J.D. and Schnell, N.F.
 Phosphomevalonate kinase (pmk) gene (eig8) from candida albicans
 Patent: WO 0114533-A 6 01-MAR-2001;
 Astrazeneca AB (SE)
 FEATURES
 source
 1..1299
 Location/Qualifiers
 /organism="Candida albicans"
 /mol_type="unassigned DNA"
 /db_xref="taxon:5476"

Query Match 100.0%; Score 1299; DB 6; Length 1399;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTTAGTGCACCTGGAAAGCATTTCTGCTGGTGGATATTTGGTCTT 60
 DB 1 ATGTCAAAAGCATTTAGTGCACCTGGAAAGCATTTCTGCTGGTGGATATTTGGTCTT 60
 QY 61 GAGCCATTATATGATGCTTATGATGACAGCATTTGATCAGAAATGATGAGTTATTAAC 120
 DB 61 GAGCCATTATATGATGCTTATGATGACAGCATTTGATCAGAAATGATGAGTTATTAAC 120
 QY 121 CCAAGAGACCACTTTGAAAGATCTAGATCAAAATTTCTTCCACCCCAATTTGCAAAAC 180

DB 121 CCAAGAGACCACTTTGAAAGATCTAGATCAAAATTTCTTCCACCCCAATTTGCAAAAC 180
 QY 181 GAGGAATGGGAATATCATATCATATCAAAATTCAGAGAGCCGAGAGAAATTCAGTCAAGC 240
 DB 181 GAGGAATGGGAATATCATATCATATCAAAATTCAGAGAGCCGAGAGAAATTCAGTCAAGC 240
 QY 241 ATAAATCCATTTTGAAGCACTATATTCATCTGCTTATGCTTATTTATTTCAACCGAA 300
 DB 241 ATAAATCCATTTTGAAGCACTATATTCATCTGCTTATTTATTTCAACCGAA 300
 QY 301 GATTTGATCTTGAATATATATATTTACTAGACCTCTGATATCTTCAAGAGATATCT 360
 DB 301 GATTTGATCTTGAATATATATATTTACTAGACCTCTGATATCTTCAAGAGATATCT 360
 QY 361 GAAACCAAGACATCTCGAATGAGAGAAAAAATTTCTTTTACCTTCTGCTGCAATACC 420
 DB 361 GAAACCAAGACATCTCGAATGAGAGAAAAAATTTCTTTTACCTTCTGCTGCAATACC 420
 QY 421 GAAATGGAAAAGCCGATTTAGTTTCATGCGCAGAGATTAGTGTAGTTGCAAGT 480
 DB 421 GAAATGGAAAAGCCGATTTAGTTTCATGCGCAGAGATTAGTGTAGTTGCAAGT 480
 QY 481 TTATATCCCATTTTATCCCAATGTTATCAGTACGAAATTAAGATTTTGGCAACGTT 540
 DB 481 TTATATCCCATTTTATCCCAATGTTATCAGTACGAAATTAAGATTTTGGCAACGTT 540
 QY 541 GCACAGATTCACATTTGTTATGCCAAAAAAGATAGATCTGCTGTTGATGTTGCAACT 600
 DB 541 GCACAGATTCACATTTGTTATGCCAAAAAAGATAGATCTGCTGTTGATGTTGCAACT 600
 QY 601 GCATTTATGCTGATTTGATTTATAGAAATTTGACCCAGCTTTGATTAATGACGTTT 660
 DB 601 GCATTTATGCTGATTTGATTTATAGAAATTTGACCCAGCTTTGATTAATGACGTTT 660
 QY 661 CAGGTTCTGAAAAGTATCTGAGAAATTCCTGAGAAATTCCTGAGAAATTTGATTAAGT 720
 DB 661 CAGGTTCTGAAAAGTATCTGAGAAATTCCTGAGAAATTCCTGAGAAATTTGATTAAGT 720
 QY 721 AACTGGGAATTCAAATGAAATGATGATACATTCATTCAGGATCAAGTATTAATGGGT 780
 DB 721 AACTGGGAATTCAAATGAAATGATGATACATTCATTCAGGATCAAGTATTAATGGGT 780
 QY 781 GACGTCAAGGTGCTCAGAAACACCCAAATGTTATGATACAGGTATCTCAATGAAAAAG 840
 DB 781 GACGTCAAGGTGCTCAGAAACACCCAAATGTTATGATACAGGTATCTCAATGAAAAAG 840
 QY 841 GAAAGCCGAAAGAAAGCTCTGTTGATGATGACAGCTTAAATGATGCAATTTACAGTTT 900
 DB 841 GAAAGCCGAAAGAAAGCTCTGTTGATGATGACAGCTTAAATGATGCAATTTACAGTTT 900
 QY 901 ATGAAGGAATTAAGGGAATATGCTGAAATAATGATGATGATGATGATGATGATGATGAT 960
 DB 901 ATGAAGGAATTAAGGGAATATGCTGAAATAATGATGATGATGATGATGATGATGATGAT 960
 QY 961 GAGTTAGATCATTTCTGTTGAGCTTTGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 961 GAGTTAGATCATTTCTGTTGAGCTTTGATGATGATGATGATGATGATGATGATGATGAT 1020
 QY 1021 CAAGCATTAACCAAAATTCAGAGTTTCAATTTGAATGATGATGATGATGATGATGATGAT 1080
 DB 1021 CAAGCATTAACCAAAATTCAGAGTTTCAATTTGAATGATGATGATGATGATGATGATGAT 1080
 QY 1081 GACCGTTGCAAGAGATTCCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 1081 GACCGTTGCAAGAGATTCCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 1141 GATGCAATAGCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1141 GATGCAATAGCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1201 GAAATCAGATTTATTTTCAATGATTTATGATGATGATGATGATGATGATGATGATGATGAT 1260

[illegible]

Qy	601	GCATTTATATGCTCATTTGATATATAGAAAGATTTCAGCAGCTTTGATATTAATGACGTGTT	660
Db	689	GCAATTTATATGCTCATTTGATATATAGAAAGATTTCAGCAGCTTTGATATTAATGACGTGTT	748
Qy	661	CAGGTTCTAGAAAGTGAATCTTGAGAAAGTTCCTCCACAGAGTTGAAAAAATTGATTGAAAGT	720
Db	749	CAGGTTCTAGAAAGTGAATCTTGAGAAAGTTCCTCCACAGAGTTGAAAAAATTGATTGAAAGT	808
Qy	721	AACGTGGGAATTCAAACATGTAAAGATGTACATTAACGATACGGAATCAAGTTATTTAATGGT	780
Db	809	AACGTGGGAATTCAAACATGTAAAGATGTACATTAACGATACGGAATCAAGTTATTTAATGGT	868
Qy	781	GACGTCAAAGGCTGCTCAGAAACACCCAAATTTGGTATCAGAGACTCCCAATGAAAAAG	840
Db	869	GACGTCAAAGGCTGCTCAGAAACACCCAAATTTGGTATCAGAGACTCCCAATGAAAAAG	928
Qy	841	GAAGAACCCAGAGAAAGCTCTGTTGTGTATGACCAAGCTTAATAGTCCCAATTTACAGTT	900
Db	929	GAAGAACCCAGAGAAAGCTCTGTTGTGTATGACCAAGCTTAATAGTCCCAATTTACAGTT	988
Qy	901	ATGAAGGAATTTGAGAGGAATGCGGTGAAAAATATGACATCAGACCAGAGACTTAATATTA	960
Db	989	ATGAAGGAATTTGAGAGGAATGCGGTGAAAAATATGACATCAGACCAGAGACTTAATATTA	1048
Qy	961	GAGTTAGATCATTTCTGTGAGACCTTTGACTGTGGCATTAAGAAATCATCAGAAAAAGGTTA	1020
Db	1049	GAGTTAGATCATTTCTGTGAGACCTTTGACTGTGGCATTAAGAAATCATCAGAAAAAGGTTA	1108
Qy	1021	CAAGCATTAACACAAAAATCAGAGGTTCCAAATTTGAACTGATGTGCCAAACCCAGTTGTG	1080
Db	1109	CAAGCATTAACACAAAAATCAGAGGTTCCAAATTTGAACTGATGTGCCAAACCCAGTTGTG	1168
Qy	1081	GACCGTTGTCAAGAGATTCCTGTTGTGTGTGGTGATGCTGTCAGAGTGCCTGTGATAC	1140
Db	1169	GACCGTTGTCAAGAGATTCCTGTTGTGTGTGGTGATGCTGTCAGAGTGCCTGTGATAC	1228
Qy	1141	GATGCAATAGCTGTATTATGTTGTGGAAAAATCAGTGGAAAAATTTTAAGCAAAAACTCTT	1200
Db	1229	GATGCAATAGCTGTATTATGTTGTGGAAAAATCAGTGGAAAAATTTTAAGCAAAAACTCTT	1288
Qy	1201	GAAGAAATCAGATTATTTCTCAATATAGTTTACTGGGTTGATTTGGAAAGCAAAACGAAGT	1260
Db	1289	GAAGAAATCAGATTATTTCTCAATATAGTTTACTGGGTTGATTTGGAAAGCAAAACGAAGT	1348
Qy	1261	GTACTTGAGAAAAAACCCAGAAAGCTATATATAGTTTATTA	1299
Db	1349	GTACTTGAGAAAAAACCCAGAAAGCTATATATAGTTTATTA	1387
RESULT 3			
AX240804			
LOCUS	AX240804	1299 bp	DNA linear PAT 26-SEP-2001
DEFINITION	Sequence 42 from Patent WO0160975.		
ACCESSION	AX240804		
VERSION	AX240804.1	GI:15797740	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 Roemer, T., Jiang, B., Boone, C. and Bussey, H.		
TITLE	Gene disruption methodologies for drug target discovery		
JOURNAL	Patent: WO 0160975-A 42 23-AUG-2001;		
FEATURES	Elittra Pharmaceuticals, Inc. (US)		
source	location/Qualifiers		
	1..1299	/organism="Candida albicans"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:5476"	

Query Match 68.6%; Score 891; DB 6; Length 1299;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 1 ATGTCAAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGCTGTGTGATATTTGGTCTT 60
DB 1 ATGTCAAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGCTGTGTGATATTTGGTCTT 60
QY 61 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGATCAGCAATGATGATGATGATTAACA 120
DB 61 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGATCAGCAATGATGATGATGATTAACA 120
QY 121 CCAAAAAGAACCAAGTTGAAAAGAAATCTGAATCAAAATTTCTTCAACCCCAATTTGCAAC 180
DB 121 CCAAAAAGAACCAAGTTGAAAAGAAATCTGAATCAAAATTTCTTCAACCCCAATTTGCAAC 180
QY 181 GGAGATGGGAATCTCATATCATCAAAATACAGAGAGCCCGAGAGAGTTCACTACACGC 240
DB 181 GGAGATGGGAATCTCATATCATCAAAATACAGAGAGCCCGAGAGAGTTCACTACACGC 240
QY 241 ATAAATCCATTTTATAGAGGCACTATATTCATGCTTTATGCTTATATTCACACCGCAA 300
DB 241 ATAAATCCATTTTATAGAGGCACTATATTCATGCTTTATGCTTATATTCACACCGCAA 300
QY 301 GCATTTGATCTGAAAATCATATTTACTGACACCTTGATATCATTTCAAGAGATTAAT 360
DB 301 GCATTTGATCTGAAAATCATATTTACTGACACCTTGATATCATTTCAAGAGATTAAT 360
QY 361 GAAACCAAGACATCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCT 420
DB 361 GAAACCAAGACATCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCT 420
QY 421 GAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 GAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 TTATATCCCATTTTATCCCAATGTTTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 TTATATCCCATTTTATCCCAATGTTTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GCACAGATTCACATTTGTTATGCTCCCAAAAAGATAGAGATCTGGGTTGATGTCACACT 600
DB 541 GCACAGATTCACATTTGTTATGCTCCCAAAAAGATAGAGATCTGGGTTGATGTCACACT 600
QY 601 GCAATTTATGCTGTGATGTTATGAGAGATTTGACAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 GCAATTTATGCTGTGATGTTATGAGAGATTTGACAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 CAGGTTCTAGAAAGTGATCTCGAGAAAGTTCCCAAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 CAGGTTCTAGAAAGTGATCTCGAGAAAGTTCCCAAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 AACTGGAATTCAAACATGAAAGATGATCATTAATCGAATCAAGTTATTAATGAGGT 780
DB 721 AACTGGAATTCAAACATGAAAGATGATCATTAATCGAATCAAGTTATTAATGAGGT 780
QY 781 GACGTCAGAGGAGTGTGCTGAGAAACCCCAATTTGATACAGAGATCTCAATGAGAGAG 840
DB 781 GACGTCAGAGGAGTGTGCTGAGAAACCCCAATTTGATACAGAGATCTCAATGAGAGAG 840
QY 841 GAAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 GAAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 ATGAAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 ATGAAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 GAGTTAGATCATTTCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 961 GAGTTAGATCATTTCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 CAACGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080

```

```

DB 1021 CAACGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 GACCGTTGCAAGAGATTCCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 GACCGTTGCAAGAGATTCCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 GATGCAATAGCTGTATATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 GATGCAATAGCTGTATATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GAAAATCCAGATTTATTTTATTAATGTTTACTGAGTTGATTTGAGAGAGAGAGAGAG 1260
DB 1201 GAAAATCCAGATTTATTTTATTAATGTTTACTGAGTTGATTTGAGAGAGAGAGAGAG 1260
QY 1261 GTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1299
DB 1261 GTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1299

```

RESULT 4
 AX488723 1299 bp DNA linear PAT 16-AUG-2002
 LOCUS AX488723
 DEFINITION Sequence 6023 from Patent WO02053728.
 ACCESSION AX488723
 VERSION AX488723.1 GI:22322735
 KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K. L.
 Gene disruption methodologies for drug target discovery
 Patent: WO 02053728-A 6023 11-JUL-2002;
 JOURNAL Elitira Pharmaceuticals, Inc. (US)
 location/Qualifiers

FEATURES

source

1..1299
 /organism="Candida albicans"
 /mol_type="unassigned DNA"
 /db_xref="taxon:5476"

ORIGIN

```

Query Match 68.6%; Score 891; DB 6; Length 1299;  

Best Local Similarity 99.4%; Pred. No. 0;  

Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGTCAAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGCTGTGTGATATTTGGTCTT 60
DB 1 ATGTCAAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGCTGTGTGATATTTGGTCTT 60
QY 61 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGATCAGCAATGATGATGATGATTAACA 120
DB 61 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGATCAGCAATGATGATGATGATTAACA 120
QY 121 CCAAAAAGAACCAAGTTGAAAAGAAATCTGAATCAAAATTTCTTCAACCCCAATTTGCAAC 180
DB 121 CCAAAAAGAACCAAGTTGAAAAGAAATCTGAATCAAAATTTCTTCAACCCCAATTTGCAAC 180
QY 181 GGAGATGGGAATCTCATATCATCAAAATACAGAGAGCCCGAGAGAGTTCACTACACGC 240
DB 181 GGAGATGGGAATCTCATATCATCAAAATACAGAGAGCCCGAGAGAGTTCACTACACGC 240
QY 241 ATAAATCCATTTTATAGAGGCACTATATTCATGCTTTATGCTTATATTCACACCGCAA 300
DB 241 ATAAATCCATTTTATAGAGGCACTATATTCATGCTTTATGCTTATATTCACACCGCAA 300
QY 301 GCATTTGATCTGAAAATCATATTTACTGACACCTTGATATCATTTCAAGAGATTAAT 360
DB 301 GCATTTGATCTGAAAATCATATTTACTGACACCTTGATATCATTTCAAGAGATTAAT 360
QY 361 GAAACCAAGACATCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCT 420

```


Query Match 25.7%; Score 334; DB 6; Length 577;
 Best Local Similarity 99.7%; Pred. No. 2.4e-163;
 Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 331 GACCCGTGATATGATTCAGCAAGAAAGATACGAAACCAAGCATCTCGAATGAGAAAA 390
 DB 421 GACCCGTGATATGATTCAGCAAGAAAGATACGAAACCAAGCATCTCGAATGAGAAAA 362
 QY 391 ACAATTTCTTTACCATTCCTCGTCGCAATTCAGAAAGTGGAAAAAGCCGGAATTAAGTTGTCG 450
 DB 361 ACTTTCTTTACCATTCCTCGTCGCAATTCAGAAAGTGGAAAAAGCCGGAATTAAGTTGTCG 302
 QY 451 GCAGGATTAAGTACGATTCGTCGCAAGATTTTATATCCCATTTTATCCCAATGTTATC 510
 DB 301 GCAGGATTAAGTACGATTCGTCGCAAGATTTTATATCCCATTTTATCCCAATGTTATC 242
 QY 511 AGTACGAATTAAGATATTTTGGCAAGCTGGCAAGATTTGACATGTTATGCCCCAAAA 570
 DB 241 AGTACGAATTAAGATATTTTGGCAAGCTGGCAAGATTTGACATGTTATGCCCCAAAA 182
 QY 571 AGATATGATCTGGGTTTGAATGTTGCAACTGCAATTTATGCTGATTTATATAGAGA 630
 DB 181 AGATATGATCTGGGTTTGAATGTTGCAACTGCAATTTATGCTGATTTATATAGAGA 122
 QY 631 TTTACGCCAGCTTTGATTAATAGACGTTTTCAGTTCTGAGAAAGTATCTTGAAGATTC 690
 DB 121 TTTACGCCAGCTTTGATTAATAGACGTTTTCAGTTCTGAGAAAGTATCTTGAAGATTC 62
 QY 691 CCCACAGAGTTGAAAAAATTTGATTG 715
 DB 61 CCCACAGAGTTGAAAAAATTTGATTG 37

RESULT 7
 LOCUS AX485723/c 65 bp DNA linear PAT 16-AUG-2002
 DEFINITION Sequence 3023 from Patent WO02053728.
 ACCESSION AX485723
 VERSION AX485723.1 GI:22319939
 KEYWORDS

SOURCE Candida albicans
 ORGANISM Candida albicans
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; mitosporic Saccharomycetales; Candida.
 REFERENCE 1 Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K.L.
 TITLE Gene disruption methodologies for drug target discovery
 JOURNAL Patent: WO 02053728-A 3023 11-JUL-2002;
 Eiltra Pharmaceuticals, Inc. (US)

FEATURES
 source
 1..65
 Location/Qualifiers
 /organism="Candida albicans"
 /mol_type="unassigned DNA"
 /db_xref="taxon:5476"

ORIGIN

Query Match 5.0%; Score 65; DB 6; Length 65;
 Best Local Similarity 100.0%; Pred. No. 3.6e-22;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTTAGTGACCTGGAAAGCATTTCTGCTGGATATTTGGTTCTT 60
 DB 65 ATGTCAAAAGCATTTAGTGACCTGGAAAGCATTTCTGCTGGATATTTGGTTCTT 6
 QY 61 GAGCC 65
 DB 5 GAGCC 1

RESULT 8
 LOCUS AX241110/c 90 bp DNA linear PAT 26-SEP-2001
 DEFINITION Sequence 348 from Patent WO0160975.

ACCESSION AX241110
 VERSION AX241110.1 GI:15797985
 KEYWORDS

SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1 Roemer, T., Jiang, B., Boone, C. and Bussey, H.
 TITLE Gene disruption methodologies for drug target discovery
 JOURNAL Patent: WO 0160975-A 348 23-AUG-2001;
 Eiltra Pharmaceuticals, Inc. (US)

FEATURES
 source
 1..90
 Location/Qualifiers
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="DNA primer"

ORIGIN

Query Match 5.0%; Score 65; DB 6; Length 90;
 Best Local Similarity 100.0%; Pred. No. 3.6e-22;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTTAGTGACCTGGAAAGCATTTCTGCTGGATATTTGGTTCTT 60
 DB 65 ATGTCAAAAGCATTTAGTGACCTGGAAAGCATTTCTGCTGGATATTTGGTTCTT 6
 QY 61 GAGCC 65
 DB 5 GAGCC 1

RESULT 9
 LOCUS AX087877/c 25 bp DNA linear PAT 17-MAR-2001
 DEFINITION Sequence 4 from Patent WO0114533.
 ACCESSION AX087877
 VERSION AX087877.1 GI:13396870
 KEYWORDS

SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1 Rosamond, J.D. and Schnell, N.F.
 TITLE Phosphomemalonate kinase (pmk) gene (erg8) from candida albicans
 JOURNAL Patent: WO 0114533-A 4 01-MAR-2001;
 Astrazeneca AB (SE)

FEATURES
 source
 1..25
 Location/Qualifiers
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Single-stranded oligonucleotide"

ORIGIN

Query Match 1.9%; Score 25; DB 6; Length 25;
 Best Local Similarity 100.0%; Pred. No. 0.37;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 477 AAGTTATATATCCCATTTATATCCCC 501
 DB 25 AAGTTATATATCCCATTTATATCCCC 1

RESULT 10
 LOCUS AX087884/c 36 bp DNA linear PAT 17-MAR-2001
 DEFINITION Sequence 11 from Patent WO0114533.
 ACCESSION AX087884
 VERSION AX087884.1 GI:13396876
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct

```

artificial sequences.
REFERENCE
1
AUTHORS
1
TITLE
Rosamond J.D. and Schnell N.F.
JOURNAL
Phosphomevalonate kinase (pmk) gene (erg8) from candida albicans
Patent: WO 0114533-A 11 01-MAR-2001;
Astrazeneca AB (SE)
FEATURES
source
1..36
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Single-stranded oligonucleotide"
ORIGIN
Query Match 1.8%; Score 24; DB 6; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1276 CCAGAGACTATAGTTTATTA 1299
|||||
36 CCAGAGACTATAGTTTATTA 13
RESULT 11
AE014174.2/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AE014174 Accession AE014174
Fragment Name Begin End
AE014174_0 1 110000
AE014174_1 100001 210000
AE014174_2 200001 310000
AE014174_3 300001 400029
Continuation (3 of 4) of AE014174 from base 200001 (AE014174 Mus musculus plebeald delcti
Query Match 1.8%; Score 23; DB 10; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1211 ATTATTTCTATGTTTACTG 1233
|||||
92633 ATTATTTCTATGTTTACTG 92611
RESULT 12
AC115701 201532 bp DNA linear HTG 09-MAR-2004
LOCUS
Mus musculus chromosome 14 clone RP23-434P22 map 14, WORKING DRAFT
SEQUENCE, 9 ordered pieces.
AC115701
AC115701.4 GI:44886761
HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Mus musculus
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 201532)
REFERENCE
1
AUTHORS
Birren, B., Nuebaum, C. and Lander, E.
JOURNAL
Unpublished
2 (bases 1 to 201532)
Birren, B., Linton, L., Nuebaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barin, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Bouhagalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, T.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Galagan, J., Gardyna, S.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatae, A., Kells, C., Labocque, K., Lamazares, R.,
Lander, E., Lenocsky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Menus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,

```

```

Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rie, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Straus, N., Subramanian, A., Talama, D., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.T., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 201532)
REFERENCE
1
AUTHORS
Birren, B., Nuebaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arabchi, H.M., Barin, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Bouhagalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatae, A., Kells, C., Lander, E., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rie, C., Rogov, P.,
Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talama, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (09-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2004 this sequence version replaced gi:31455673.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20308
Center clone name: 434_P 22
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 197912 bases at least Q40
Consensus quality: 198811 bases at least Q30
Consensus quality: 199046 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 199227; sum-of-ctngs
Quality coverage: 10.9 in Q20 bases; agarose-fp
Quality coverage: 10.9 in Q20 bases; sum-of-ctngs
-----
NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 92555: contig of 92555 bp in length
92556 92555: gap of 100 bp
92556 92555: gap of 100 bp
92556 92555: contig of 2745 bp in length
92556 92555: gap of 100 bp

```

```

* 95501 130686: contig of 35186 bp in length
* 130687 130786: gap of 100 bp
* 130787 134173: contig of 3387 bp in length
* 134174 134273: gap of 100 bp
* 134274 166265: contig of 31992 bp in length
* 166266 166365: gap of 100 bp
* 166366 170490: contig of 4125 bp in length
* 170491 170590: gap of 100 bp
* 170591 181579: contig of 10989 bp in length
* 181580 181679: gap of 100 bp
* 181680 183883: contig of 2204 bp in length
* 183884 183983: gap of 100 bp
* 183984 201532: contig of 17549 bp in length.
FEATURES
source
1..201532
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="14"
/map="14"
/clone="RP23-434P22"
/clone_id="RPC1-23 Female Mouse BAC"
1..92555
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
92556..95400
/note="assembly_fragment"
95501..130686
/note="assembly_fragment"
130787..134173
/note="assembly_fragment"
134274..166265
/note="assembly_fragment"
166366..170490
/note="assembly_fragment"
170591..181579
/note="assembly_fragment"
181680..183883
/note="assembly_fragment"
183984..201532
/note="assembly_fragment"
clone_end:T7
vector_side:right"

```

ORIGIN

Query Match 1.8%; Score 23; DB 2; Length 201532;
 Best Local Similarity 100.0%; Pred. No. 2.5;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1211 ATTATTTTCATATGTTTACTG3 1233
 Db 44001 ATTATTTTCATATGTTTACTG3 44023

RESULT 13
 AX241232/c
 LOCUS AX241232 22 bp DNA linear PAT 26-SEP-2001
 DEFINITION Sequence 470 from Patent WO0160975.
 ACCESSION AX241232
 VERSION AX241232.1 GI:15798107
 KEYWORDS
 SOURCE .
 ORGANISM synthetic construct
 artificial sequences.
 REFERENCE 1
 AUTHORS Roemer,T., Jiang,B., Boone,C. and Bussey,H.
 TITLE Gene disruption methodologies for drug target discovery
 JOURNAL Patent: WO 0160975-A 470 23-AUG-2001;
 Eiltra Pharmaceuticals, Inc. (US)
 FEATURES
 source
 1..22
 /organism="synthetic construct"

ORIGIN

```

/mol_type="unassigned DNA"
/db_xref="taxon:33630"
/note="DNA primer"

```

Query Match 1.7%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 454 GGATTAGTGTCAAGTTGTGCCA 475
 Db 22 GGATTAGTGTCAAGTTGTGCCA 1

RESULT 14
 AX487723/c
 LOCUS AX487723 22 bp DNA linear PAT 16-AUG-2002
 DEFINITION Sequence 5023 from Patent WO02051728.
 ACCESSION AX487723
 VERSION AX487723.1 GI:22321803
 KEYWORDS
 SOURCE .
 ORGANISM Candida albicans
 Candida albicans
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; mitosporic Saccharomycetales; Candida.
 REFERENCE 1
 AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
 TITLE Gene disruption methodologies for drug target discovery
 JOURNAL Patent: WO 02051728-A 5023 11-JUL-2002;
 Eiltra Pharmaceuticals, Inc. (US)
 FEATURES
 source
 1..22
 /organism="Candida albicans"
 /mol_type="unassigned DNA"
 /db_xref="taxon:5476"

ORIGIN

Query Match 1.7%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 454 GGATTAGTGTCAAGTTGTGCCA 475
 Db 22 GGATTAGTGTCAAGTTGTGCCA 1

RESULT 15
 AC090350
 LOCUS AC090350 58879 bp DNA linear HTG 15-APR-2001
 DEFINITION Homo sapiens chromosome 11 clone RP11-655N17 map 11, LOW-PASS
 SEQUENCE SAMPLING.
 ACCESSION AC090350
 VERSION AC090350.3 GI:13625484
 KEYWORDS HTG: HTGS PHASE0
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 58879)
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE Homo sapiens chromosome 11, clone RP11-655N17
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 58879)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
 Batra,N., Baatien,V., Boguslavsky,L., Bouckgalter,B., Brown,A.,
 Camarata,J., Campopiano,A., Choedel,Y., Colangelo,M., Collins,S.,
 Collumore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
 Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Karates,A., Lacroque,K., Lamazares,R., Landers,T.,
 Lenocksky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
 Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,

TITLE
JOURNAL
COMMENT

McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norb, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roettli, M.,
Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Souarez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strasser, N., Subramanian, A., Talama, J., Tesfaye, S., Theodore, J.,
Travis, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J.,
Zemek, L., Zimmer, A. and Zody, W.

Direct Submission

Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 15, 2001 this sequence version replaced gi:13431044.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L12631

Center clone name: 655_N_17

* NOTE: This record contains 70 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 760: contig of 760 bp in length
* 761 860: gap of 100 bp
* 861 1563: contig of 703 bp in length
* 1564 1663: gap of 100 bp
* 1664 2379: contig of 716 bp in length
* 2380 2479: gap of 100 bp
* 2480 3209: contig of 730 bp in length
* 3210 3309: gap of 100 bp
* 3310 4045: contig of 736 bp in length
* 4046 4145: gap of 100 bp
* 4146 4887: contig of 742 bp in length
* 4888 4987: gap of 100 bp
* 4988 5757: contig of 770 bp in length
* 5758 5857: gap of 100 bp
* 5858 6609: contig of 752 bp in length
* 6610 6709: gap of 100 bp
* 6710 7434: contig of 725 bp in length
* 7435 7534: gap of 100 bp
* 7535 8289: contig of 755 bp in length
* 8290 8390: gap of 100 bp
* 8390 9138: contig of 749 bp in length
* 9139 9238: gap of 100 bp
* 9239 9985: contig of 747 bp in length
* 9986 10085: gap of 100 bp
* 10086 10823: contig of 738 bp in length
* 10824 10923: gap of 100 bp
* 10924 11672: contig of 749 bp in length
* 11673 11772: gap of 100 bp
* 11773 12501: contig of 729 bp in length
* 12502 12601: gap of 100 bp
* 12602 13350: contig of 749 bp in length
* 13351 13450: gap of 100 bp
* 13451 14181: contig of 731 bp in length
* 14182 14281: gap of 100 bp
* 14282 15006: contig of 725 bp in length

15007 15106: gap of 100 bp
* 15107 15842: contig of 736 bp in length
* 15843 15942: gap of 100 bp
* 15943 16594: contig of 752 bp in length
* 16595 16794: gap of 100 bp
* 16795 17547: contig of 753 bp in length
* 17548 17647: gap of 100 bp
* 17648 18396: contig of 749 bp in length
* 18397 18496: gap of 100 bp
* 18497 19235: contig of 739 bp in length
* 19236 19335: gap of 100 bp
* 19336 20099: contig of 764 bp in length
* 20100 20199: gap of 100 bp
* 20200 20945: contig of 746 bp in length
* 20946 21045: gap of 100 bp
* 21046 21794: contig of 749 bp in length
* 21795 21895: gap of 100 bp
* 21895 22622: contig of 728 bp in length
* 22623 22722: gap of 100 bp
* 22723 23451: contig of 729 bp in length
* 23452 23551: gap of 100 bp
* 23552 24286: contig of 735 bp in length
* 24287 24386: gap of 100 bp
* 24387 25122: contig of 736 bp in length
* 25123 25222: gap of 100 bp
* 25223 25976: contig of 754 bp in length
* 25977 26076: gap of 100 bp
* 26077 26827: contig of 751 bp in length
* 26828 26927: gap of 100 bp
* 26928 27676: contig of 749 bp in length
* 27677 28522: contig of 746 bp in length
* 28523 28622: gap of 100 bp
* 28623 29373: contig of 751 bp in length
* 29374 29473: gap of 100 bp
* 29474 30268: contig of 795 bp in length
* 30269 31114: contig of 746 bp in length
* 31115 31214: gap of 100 bp
* 31215 31939: contig of 725 bp in length
* 31940 32039: gap of 100 bp
* 32040 32771: contig of 732 bp in length
* 32772 32871: gap of 100 bp
* 32872 33601: contig of 730 bp in length
* 33602 33701: gap of 100 bp
* 33702 34435: contig of 734 bp in length
* 34436 34535: gap of 100 bp
* 34536 35281: contig of 746 bp in length
* 35282 35381: gap of 100 bp
* 35382 36120: contig of 735 bp in length
* 36121 36220: gap of 100 bp
* 36221 36966: contig of 746 bp in length
* 36967 37066: gap of 100 bp
* 37067 37797: contig of 731 bp in length
* 37798 37897: gap of 100 bp
* 37898 38641: contig of 744 bp in length
* 38642 38741: gap of 100 bp
* 38742 39526: contig of 785 bp in length
* 39527 39626: gap of 100 bp
* 39627 40465: contig of 739 bp in length
* 40466 41303: gap of 100 bp
* 41304 41303: contig of 738 bp in length
* 41304 42041: contig of 738 bp in length
* 42042 42141: gap of 100 bp
* 42142 42891: contig of 750 bp in length
* 42892 42991: gap of 100 bp
* 42992 43728: contig of 737 bp in length
* 43729 43828: gap of 100 bp
* 43829 44572: contig of 744 bp in length
* 44573 44672: gap of 100 bp
* 44673 45412: contig of 740 bp in length
* 45413 45512: gap of 100 bp

* 45513 46227: contig of 715 bp in length
* 46228 46327: gap of 100 bp
* 46328 47081: contig of 754 bp in length
* 47082 47181: gap of 100 bp
* 47182 47934: contig of 753 bp in length
* 47935 48034: gap of 100 bp
* 48035 48790: contig of 756 bp in length
* 48791 48890: gap of 100 bp
* 48891 49630: contig of 740 bp in length
* 49631 49730: gap of 100 bp
* 49731 50471: contig of 741 bp in length
* 50472 50571: gap of 100 bp
* 50572 51316: contig of 745 bp in length
* 51317 51416: gap of 100 bp
* 51417 52149: contig of 733 bp in length
* 52150 52249: gap of 100 bp
* 52250 52995: contig of 746 bp in length
* 52996 53095: gap of 100 bp
* 53096 53845: contig of 750 bp in length
* 53846 53945: gap of 100 bp
* 53946 54673: gap of 727 bp in length
* 54673 55221: contig of 749 bp in length
* 55222 55622: gap of 100 bp
* 55622 56364: contig of 743 bp in length
* 56365 56464: gap of 100 bp
* 56465 57190: contig of 726 bp in length
* 57191 57290: gap of 100 bp
* 57291 58033: contig of 743 bp in length

Query Match 1.7%; Score 22; DB 2; length 58879;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yr 1030 ACACAAATCATGAGCTTCCA 1051
Db 19543 ACACAAATCATGAGCTTCCA 19564

RESULT 16 AC149142 LOCUS

DEFINITION Xenopus tropicalis clone CH216-82H3, *** SEQUENCING IN PROGRESS
AC149142 153598 bp DNA linear HTG 20-MAY-2004
AC149142
AC149142.2 GI:47523990
VERSION HTG; HTGS PHASRI.
KEYWORDS Xenopus tropicalis (Silurana tropicalis)
SOURCE Xenopus tropicalis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 153598)
DOE Joint Genome Institute.
Unpublished
2 (bases 1 to 153598)
DOE Joint Genome Institute.
Direct Submission
Submitted (19-MAY-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
94598-1698, USA
3 (bases 1 to 153598)
DOE Joint Genome Institute.
Direct Submission
Submitted (20-MAY-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
94598-1698, USA
On May 20, 2004 this sequence version replaced gi:47498095.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 3599316
Center clone name: CHOR-216_82H3

Summary Statistics
Consensus quality: 141761 bases at least Q40
Consensus quality: 145817 bases at least Q30
Consensus quality: 148968 bases at least Q20
Estimated insert size: 170000; agarose-ff estimation
Estimated insert size: 150998; sum-of-contigs estimation
Quality coverage: 2.42 in Q20 bases; agarose-ff estimation
Quality coverage: 2.72 in Q20 bases; sum-of-contigs estimation
NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2637 2636: contig of 2636 bp in length
* 2637 2736: gap of unknown length
* 2737 4668: contig of 1932 bp in length
* 4669 4768: gap of unknown length
* 4769 6660: contig of 1892 bp in length
* 6661 6760: gap of unknown length
* 6761 8966: contig of 2206 bp in length
* 8967 9066: gap of unknown length
* 9067 12223: contig of 3157 bp in length
* 12224 12323: gap of unknown length
* 12324 14853: contig of 2530 bp in length
* 14854 14953: gap of unknown length
* 14954 17543: contig of 2590 bp in length
* 17544 17643: gap of unknown length
* 17644 19766: contig of 2123 bp in length
* 19767 19866: gap of unknown length
* 19867 21772: contig of 1906 bp in length
* 21773 21872: gap of unknown length
* 21873 25494: contig of 3622 bp in length
* 25495 25594: gap of unknown length
* 25595 28992: contig of 3398 bp in length
* 28993 31780: gap of unknown length
* 31781 31880: contig of 2688 bp in length
* 31881 36776: gap of unknown length
* 36777 36876: gap of unknown length
* 36877 39602: contig of 2726 bp in length
* 39603 39702: gap of unknown length
* 39703 44268: contig of 4566 bp in length
* 44269 44368: gap of unknown length
* 44369 48059: contig of 3631 bp in length
* 48060 48159: gap of unknown length
* 48160 53328: contig of 5169 bp in length
* 53329 53428: gap of unknown length
* 53429 58675: contig of 5247 bp in length
* 58676 58775: gap of unknown length
* 58776 66270: contig of 7495 bp in length
* 66271 66370: gap of unknown length
* 66371 73950: contig of 7580 bp in length
* 73951 74050: gap of unknown length
* 74051 83527: contig of 9477 bp in length
* 83528 83627: gap of unknown length
* 83628 90690: contig of 7063 bp in length
* 90691 90790: gap of unknown length
* 90791 99079: contig of 8289 bp in length
* 99080 99179: gap of unknown length
* 99180 109287: contig of 10108 bp in length
* 109288 109387: gap of unknown length
* 109388 119271: contig of 9884 bp in length
* 119272 119371: gap of unknown length
* 119372 132522: contig of 13151 bp in length
* 132523 132622: gap of unknown length

FEATURES	* 132623	153598: contig of 20976 bp in length.
SOURCE	Location/Qualifiers	
	1. 153598	
	/organism="Xenopus tropicalis"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:8364"	
	/clone="CH216-82H3"	
	/clone_lib="CHORI-216 Xenopus tropicalis BAC library"	
ORIGIN		
Query Match	1.7%; Score 22; DB 2; Length 153598;	
Best Local Similarity	100.0%; Pred. No. 8.5;	
Matches	22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	569 AAAAGATGAGATCTGGGTTTGA 590	
DB	68052 AAAAGATGAGATCTGGGTTTGA 68073	
RESULT 17		
LOCUS	AC103846 171133 bp DNA linear PRI 23-JUL-2002	
DEFINITION	Homo sapiens chromosome 8, clone RP11-645H17, complete sequence.	
ACCESSION	AC103846	
VERSION	AC103846.2 GI:21930254	
KEYWORDS	HTG.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE	1 (bases 1 to 171133)	
JOURNAL	Birren, B., Nusbaum, C. and Lander, E.	
REFERENCE	Unpublished	
AUTHORS	2 (bases 1 to 171133)	
	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,	
	Anderson, S., Barna, N., Baetien, V., Boguslavsky, L., Bouckgatter, B.,	
	Brown, A., Camarata, J., Campolillo, A., Chang, Y., Chazaro, B.,	
	Chapel, Y., Colangelo, M., Coppolino, S., Collymore, A., Cook, A.,	
	Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S.,	
	Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,	
	Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,	
	Hagos, B., Heaford, A., Horton, L., Hulme, M., Iliev, I., Johnson, R.,	
	Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,	
	Maclean, C., Macdonald, P., Major, J., Levine, R., Liu, G.,	
	McCarthy, M., McKean, P., McKernan, K., McSheeters, R., Meldrim, J.,	
	Menesh, L., Mihova, T., Mlenga, V., Murphy, T., Neylor, J., Nguyen, C.,	
	Notebu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,	
	Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,	
	Raymond, C., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,	
	Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,	
	Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,	
	Strauss, N., Subramanian, A., Talamas, J., Teeffaye, S., Theodore, J.,	
	Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,	
	Viola, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,	
	Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
	Direct Submission	
JOURNAL	Submitted (320-NOV-2001) Whitehead Institute/MIT Center for Genome	
REFERENCE	Research, 320 Charles Street, Cambridge, MA 02141, USA	
AUTHORS	3 (bases 1 to 171133)	
	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,	
	Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Bouckgatter, B.,	
	Camarata, J., Chang, Y., Chazaro, B., Chapel, Y., Collymore, A.,	
	Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,	
	Fato, S., Ferreira, P., Fitzerald, M., Gage, D., Galagan, J.,	
	Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,	
	Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamat, A.,	
	Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,	
	Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,	
	McCarthy, M., Meldrim, J., Menesh, L., Mihova, T., Mlenga, V.,	
	Murphy, T., Neylor, J., Nguyen, C., Nicol, R., Notebu, C., Norman, C.H.,	
	O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,	

TITLE	JOURNAL	COMMENT
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smtih, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Teafaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		
Submitted (23-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
On Jul 23, 2002 this sequence version replaced gi:17149731.		
All repeats were identified using RepeatMasker:		
Smt, A.F.A. & Green, P. (1996-1997)		
http://ftp.genome.washington.edu/RM/RepeatMasker.html		

Genome Center		
Center: Whitehead Institute/ MIT Center for Genome Research		
Center code: MIBR		
Web site: http://www-seq.wi.mit.edu		
Contact: sequence.submissions@genome.wi.mit.edu		

Project Information		
Center project name: L21845		
Center clone name: 645_H_17		

Location/Qualifiers		
1. 171133		
/organism="Homo sapiens"		
/mol_type="genomic DNA"		
/db_xref="taxon:9606"		
/chromosome="8"		
/map="8"		
/clone="RP11-645H17"		
/clone_lib="RPCT-11 Human Male BAC"		
complement(1008..1061)		
/rpt_family="MIR3"		
2938..3004		
/rpt_family="L2"		
complement(3035..3319)		
/rpt_family="AlusG"		
3325..3353		
/rpt_family="(T)n"		
complement(4041..4138)		
/rpt_family="L1M4"		
4139..4436		
/rpt_family="AluB"		
complement(4437..4566)		
/rpt_family="L1M4"		
complement(4587..4691)		
/rpt_family="U6"		
complement(4725..4803)		
/rpt_family="MIR3"		
5077..5307		
/rpt_family="MIR"		
5510..5804		
/rpt_family="AlusX"		
5895..5923		
/rpt_family="MIR"		
complement(7310..7475)		
/rpt_family="AlusG"		
complement(7476..7768)		
/rpt_family="AlusG"		
8818..9095		
/rpt_family="AluB"		
complement(10244..10597)		
/rpt_family="THE1C"		
11648..11960		
/rpt_family="AluY"		
12700..12826		
/rpt_family="(TG)n"		
13281..13401		
/rpt_family="MIR"		
13452..13606		
/rpt_family="MIR3"		
14111..14285		
/rpt_family="MIR"		

```

repeat_region complement(14296..14473)
/rpt_family="AluJo"
repeat_region complement(14474..14780)
/rpt_family="AluSx"
repeat_region complement(14781..14871)
/rpt_family="AluJo"
repeat_region 17534..17721
/rpt_family="MIR3"
repeat_region 18995..19277
/rpt_family="AluSx"
repeat_region 20245..20309
/rpt_family="TTTGA)n"
repeat_region complement(20328..20623)
/rpt_family="AluJo"
repeat_region 23313..23533
/rpt_family="MIR"
repeat_region complement(24402..24488)
/rpt_family="MIR"
repeat_region 24872..24946
/rpt_family="MERSA"
repeat_region 24947..25235
/rpt_family="AluJo"
repeat_region 25236..25346
/rpt_family="MERSA"
repeat_region 26121..26158
/rpt_family="TTTGA)n"
repeat_region 26818..27092
/rpt_family="AluJo"
repeat_region 27112..27353
/rpt_family="AluY"
repeat_region 27505..27525
/rpt_family="AT_rich"
repeat_region 28048..28068
/rpt_family="AT_rich"
repeat_region complement(28075..28322)
/rpt_family="HAL1"
repeat_region complement(28325..28614)
/rpt_family="HAL1"
repeat_region complement(28615..29122)
/rpt_family="MLT2B3"
repeat_region complement(29123..29234)
/rpt_family="HAL1"
repeat_region 29235..29353
/rpt_family="TMA7"
repeat_region complement(29354..29366)
/rpt_family="HAL1"
repeat_region complement(29367..29648)
/rpt_family="AluSx"
repeat_region complement(29649..30214)
/rpt_family="HAL1"
repeat_region complement(30244..30533)
/rpt_family="AluSx"
repeat_region complement(30606..30761)
/rpt_family="HAL1"
repeat_region complement(30762..30883)
/rpt_family="HAL1"
repeat_region complement(30884..31173)
/rpt_family="AluJo"
repeat_region complement(31174..31202)
/rpt_family="HAL1"
repeat_region complement(31466..31674)
/rpt_family="HAL1"
repeat_region complement(31721..31800)
/rpt_family="MADE1"
repeat_region complement(32006..32161)
/rpt_family="AluSg/x"
repeat_region 34492..34660
/rpt_family="MIR"
repeat_region 34799..34834
/rpt_family="AT_rich"
repeat_region 35099..35270
/rpt_family="MIR"
repeat_region 36132..36429

```

```

repeat_region complement(36679..36947)
/rpt_family="AluJo"
Query Match 1.7%; Score 22; DB 9; Length 171133;
Best Local Similarity 100.0%; Pred.No.8.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1030 ACACAAATCGAGGTTCCAA 1051
DB 150942 ACACAAATCGAGGTTCCAA 150921

RESULT 18
AC012311
LOCUS
DEFINITION
AC012311
AC012311.4 GI:7690196
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 185418)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 185418)
DOE Joint Genome Institute.
Direct Submission
Submitted (23-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On May 4, 2000 this sequence version replaced gi:7458754.

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----

Project Information
Center Project Name: 393685
Center clone name: CIT-HSPC_542P17

-----
Summary Statistics
Consensus quality: 171415 bases at least Q40
Consensus quality: 180780 bases at least Q30
Consensus quality: 182015 bases at least Q20
Estimated insert size: 185000; agarose-fp estimation
Estimated insert size: 184518; sum-of-contigs estimation
Quality coverage: 8.44 in Q20 bases; agarose-fp estimation
Quality coverage: 8.44 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2067: contig of 2067 bp in length
2068
2167: gap of unknown length
2168
4455: contig of 2288 bp in length
4456
4555: gap of unknown length
4556
7647: contig of 3092 bp in length
7648
7747: gap of unknown length
7748
9887: contig of 2150 bp in length
9888
9997: gap of unknown length
9998
14021: contig of 4024 bp in length
14022
25375: contig of 11254 bp in length
25376
25475: gap of unknown length
25476
48014: contig of 22539 bp in length
48015
48114: gap of unknown length

```

```

FEATURES
  source
    1. 48115 96080: contig of 47966 bp in length
    * 96081 96180: gap of unknown length
    * 96181 141589: contig of 45409 bp in length
    * 141590 141689: gap of unknown length
    * 141690 185418: contig of 43729 bp in length.
      Location/Qualifiers
        1. 185418
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"
          /chromosome="16"
          /clone="CTC-542P17"
          /clone_lib="Caltech human BAC library C"

ORIGIN
  Query Match 1.7%; Score 22; DB 2; Length 185418;
  Best Local Similarity 100.0%; Pred.No. 8.4;
  Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  32 CATTCTTCTGCTGATATTT 53
  18093 CATTCTTCTGCTGATATTT 18114

RESULT 19
AC104393/c 188270 bp DNA linear PRI 25-MAY-2002
LOCUS Homo sapiens chromosome 8, clone CTD-3080F16, complete sequence.
DEFINITION AC104393
VERSION AC104393.6 GI:21206340
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  1 (bases 1 to 188270)
  Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  Homo sapiens chromosome 8, clone CTD-3080F16
  Unpublished
  2 (bases 1 to 188270)
  Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
  Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouhgalter,B.,
  Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
  Chopel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
  Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
  Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
  Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
  Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
  Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
  Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
  Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
  McCarthy,M., McEwan,P., McKernan,K., McNeeters,R., Meldrum,J.,
  Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
  Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
  Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
  Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
  Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
  Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
  Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
  Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
  Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
  Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  Direct Submission
  Submitted (08-DEC-2001) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  3 (bases 1 to 188270)
  Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
  Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
  Bouhgalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J.,
  Chazaro,B., Chopel,Y., Colangelo,M., Collins,S., Collymore,A.,
  Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
  Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
  Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

```

```

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Labocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McNeeters,R., Meldrum,J.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  Direct Submission
  Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  4 (bases 1 to 188270)
  Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
  Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
  Bouhgalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J.,
  Chazaro,B., Chopel,Y., Colangelo,M., Collins,S., Collymore,A.,
  Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
  Faro,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D.,
  Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
  Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
  Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
  Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
  Liu,G., Maclean,C., MacDonald,P., Major,J., Margulis,N.,
  Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J.,
  Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
  Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
  O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
  Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
  Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
  Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
  Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Testaye,S.,
  Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
  Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
  Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  Direct Submission
  Submitted (25-MAY-2002) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  On May 25, 2002 this sequence version replaced gi:20429465.
  All repeats were identified using RepeatMasker:
  Smit,A.F.A. & Green,P. (1996-1997)
  http://ftp.genome.washington.edu/RM/RepeatMasker.html
  ----- Genome Center
  Center: Whitehead Institute/ MIT Center for Genome Research
  Center code: WIMR
  Web site: http://www-seq.wi.mit.edu
  Contact: sequence_submissions@genome.wi.mit.edu
  ----- Project Information
  Center project name: L22225
  Center clone name: 3080_F_16
  -----
  FEATURES
    source
      1. 188270
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="8"
        /map="8"
        /clone="CTD-3080F16"
        /clone_lib="CTD2 Human BAC"
      277. 306
        /rpt_family="AT_rich"
      repeat_region
        553_909
        /rpt_family="AluSq"
      repeat_region
        1293_1316
        /rpt_family="GC_rich"
      repeat_region
        4227_4260

```

repeat_region	32748..32822	/rpt_family="MERSA"
repeat_region	32823..33111	/rpt_family="AluIo"
repeat_region	33112..33222	/rpt_family="MERSA"
repeat_region	33997..34034	
Query Match	1.7%	Score 22; DB 9; Length 188270;
Best Local Similarity	100.0%	Pred. No. 8.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1030 ACACAAATCAGAGTCCAA 1051	
Db	158804 ACACAAATCAGAGTCCAA 158783	
RESULT 20		
AC007614/c		
LOCUS	AC007614	192347 bp DNA linear PRI 22-MAR-2003
DEFINITION	Homo sapiens chromosome 16 clone RP11-491F9, complete sequence.	
ACCESSION	AC007614	
VERSION	AC007614.7	GI:29029215
KEYWORDS	HTG.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
AUTHORS	1 (bases 1 to 192347)	
TITLE	DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.	
JOURNAL	Direct Submission	
REFERENCE	2 (bases 1 to 192347)	
AUTHORS	Jones, D., Mundt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D., Bryant, J., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L., Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R., McWhirry, K., Han, C. and Deaven, L.	
TITLE	Direct Submission	
JOURNAL	Submitted (20-MAY-1999) Center for Human Genome Studies, DOE Joint Genome Institute, Los Alamos National Laboratory, MS M888, Los Alamos, NM 87545, USA	
REFERENCE	3 (bases 1 to 192347)	
AUTHORS	DOE Joint Genome Institute.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-NOV-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	
REFERENCE	4 (bases 1 to 192347)	
AUTHORS	DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.	
TITLE	Direct Submission	
JOURNAL	Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	
COMMENT	On Mar 18, 2003 this sequence version replaced g1:16596519.	
FEATURES	Draft Sequence Produced by DOE Joint Genome Institute	
source	www.jgi.doe.gov	
source	Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory	
source	www-bhg.stanford.edu	
source	Quality: Parag Quality >=40 100% of Sequence;	
source	Estimated Total Number of Errors is 0.	
source	Location/Qualifiers	
source	1..192347	
source	/organism="Homo sapiens"	
source	/mol_type="genomic DNA"	

ORIGIN /chromosome="16"
/clone="RP11-491P9"

Query Match 1.7%; Score 22; DB 9; Length 192347;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 CATTTCTGCTGCTGATATT 53
DB 122248 CATTTCTGCTGCTGATATT 122227

RESULT 21
AX087875 21 bp DNA linear PAT 17-MAR-2001
LOCUS AX087875
DEFINITION Sequence 2 from Patent WO0114533.
ACCESSION AX087875
VERSION AX087875.1 GI:13396868
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rosemond, J.D. and Schnell, N.F.
TITLE Phosphomevalonate kinase (pmk) gene (erg8) from candida albicans
JOURNAL Patent: WO 0114533-A 2 01-MAR-2001;
Abstrazeneca AB (SE)
FEATURES
source Location/Qualifiers
1..21
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Single-stranded oligonucleotide"

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1129 GCTGTGATACGATGCAATA 1149
DB 1 GCTGTGATACGATGCAATA 21

RESULT 22
CQ661526 313 bp DNA linear PAT 03-FEB-2004
LOCUS CQ661526
DEFINITION Sequence 6452 from Patent WO02070737.
ACCESSION CQ661526
VERSION CQ661526.1 GI:42134774
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
JOURNAL liew, C.C., Marshall, W.E. and Zhang, H.
Compositions and methods relating to osteoarthritis
Patent: WO 02070737-A 6452 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source Location/Qualifiers
1..313
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 313;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGCTTCAGTCTGCTG 1135

DB 36 GTGTGCTTCAGTCTGCTG 56

RESULT 23
AX396511 479 bp DNA linear PAT 18-MAY-2002
LOCUS AX396511
DEFINITION Sequence 726 from Patent WO0212328.
ACCESSION AX396511
VERSION AX396511.1 GI:21067258
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS King, G.E., Meagher, M.J., Xu, J. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0212328-A 726 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..479
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 479;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGCTTCAGTCTGCTG 1135
DB 41 GTGTGCTTCAGTCTGCTG 61

RESULT 24
AX397039/c 530 bp DNA linear PAT 18-MAY-2002
LOCUS AX397039
DEFINITION Sequence 1254 from Patent WO0212328.
ACCESSION AX397039
VERSION AX397039.1 GI:21067786
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS King, G.E., Meagher, M.J., Xu, J. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0212328-A 1254 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..530
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 530;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGCTTCAGTCTGCTG 1135
DB 439 GTGTGCTTCAGTCTGCTG 419

RESULT 25
BD020934 771 bp DNA linear PAT 27-AUG-2002
LOCUS BD020934

```

DEFINITION Novel gene and novel gene fragment cloned in human neuroblastoma.
ACCESSION BD020934
VERSION BD020934.1 GI:22562110
KEYWORDS JP 2001245671-A/3172.
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 771)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE Novel gene and novel gene fragment cloned in human neuroblastoma
JOURNAL Patent: JP 2001245671-A 3172 11-SEP-2001;
OS CHIBA PREF. HISAMITSU PHARMACEUTICAL CO INC
PN Homo sapiens (human)
PD JP 2001245671-A/3172
PF 11-SEP-2001
PI 07-MAR-2000 JP 2000159195
PC AKIRA NAKAGAWARA
PC C12N15/09, C12Q1/68, G01N33/53, G01N33/566// (C12Q1/68, C12R1.91),
CC C12N15/00
CC Novel gene and novel gene fragment cloned in human CC
neuroblastoma
FH Key
FT source
  1. .771
  /location/Qualifiers
  1. .771
  /organism="Homo sapiens (human)".
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"

ORIGIN
Query Match 1.6%; Score 21; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1115 GTGTGTTCCAGTGTGTGTG 1135
    |||||||
Db 436 GTGTGTTCCAGTGTGTGTG 456

RESULT 26
LOCUS BD100872 771 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel genes cloned in human neuroblastoma and fragments thereof.
ACCESSION BD100872.1 GI:22646446
VERSION BD100872.1
KEYWORDS WO 0166719-A/3172.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 771)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE Novel genes cloned in human neuroblastoma and fragments thereof
JOURNAL Patent: WO 0166719-A 3172 13-SEP-2001;
OS CHIBA PREF. HISAMITSU PHARMACEUTICAL CO INC, AKIRA NAKAGAWARA
PN Homo sapiens (human)
PD WO 0166719-A/3172
PF 13-SEP-2001
PI 02-MAR-2001 WO 2001JP001629
PI 07-MAR-2000 JP 00P 159195
PC AKIRA NAKAGAWARA
PC C12N15/11, C12Q1/68, G01N33/53, G01N33/566
CC Novel genes cloned in human neuroblastoma and fragments thereof
FH Key
FT source
  1. .771
  /location/Qualifiers
  1. .771
  /organism="Homo sapiens (human)".
  /location/Qualifiers
  1. .771
  /organism="Homo sapiens (human)".
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"

FEATURES
source
  1. .771
  /organism="Homo sapiens (human)".
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"

ORIGIN
Query Match 1.6%; Score 21; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1115 GTGTGTTCCAGTGTGTGTG 1135
    |||||||
Db 436 GTGTGTTCCAGTGTGTGTG 456

RESULT 27
LOCUS CQ717294 1444 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 3228 from Patent WO02068579.
ACCESSION CQ717294
VERSION CQ717294.1 GI:42278151
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
JOURNAL Patent: WO 02068579-A 3228 06-SEP-2002;
PE Corporation (NY) (US)
FH Key
FT source
  1. .1444
  /location/Qualifiers
  1. .1444
  /organism="Homo sapiens"
  /mol_type="unassigned DNA"
  /db_xref="taxon:9606"

ORIGIN
Query Match 1.6%; Score 21; DB 6; Length 1444;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1115 GTGTGTTCCAGTGTGTGTG 1135
    |||||||
Db 1303 GTGTGTTCCAGTGTGTGTG 1323

RESULT 28
LOCUS AF068483 1596 bp mRNA linear MAM 10-JUN-1998
DEFINITION Oryctolagus cuniculus chaperonin Cct6 mRNA, complete cde.
ACCESSION AF068483
VERSION AF068483.1 GI:3201993
KEYWORDS
SOURCE Oryctolagus cuniculus (rabbit)
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 1596)
AUTHORS Schwartz, G.J., Segal, G.B. and Kittelberger, A.M.
TITLE Rabbit Chaperonin Cct6 (Tcpx)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1596)
AUTHORS Schwartz, G.J., Segal, G.B. and Kittelberger, A.M.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1998) Pediatrics, University of Rochester, 601
Elmwood Ave, Rochester, NY 14642, USA
FH Key
FT source
  1. .1596
  /location/Qualifiers
  1. .1596
  /organism="Oryctolagus cuniculus"
  /mol_type="mRNA"
  /strain="New Zealand"
  /db_xref="taxon:9986"
  /note="Tcpx"

FEATURES
source
  1. .1596
  /organism="Oryctolagus cuniculus"
  /mol_type="mRNA"
  /strain="New Zealand"
  /db_xref="taxon:9986"
  /note="Tcpx"

```

/codon_start=1
/product="chaperonin Cct6"
/protein_id="AAC19379.1"
/db_xref="GI:3201994"
/translation="MAAVKTLNPKAEVAPAAQALAVNISANGLODVLRTNLGPKGTM
KMLVSGAGDILKLTGDNVILHMOIQHPASLIARVATQDDITGDTSSNVLIIGEL
LKOADLYISEGLHPIITEGFEAAKEKALQVEIQKVSREMDRETLIDVARTSLRTKV
HAEADVLTEAVVDSILAIKKODEPIDLPEVMEKHSKSEDTSLIRGLVDHGARR
PMKRRVEDAYILTCNVSLSEYKTEVNSGFPFKSAEERKVLKARERKTFEDVKKIVE
LKKVCGSDGKGFVIVNKGIDPESDLAKGIVALKRARRNNERLTLACGVALN
SLDDINPDLGHAAGLYETLTSEKFTLEKCNPRSVTLVKGPKHTLTQIKAIR
DGLRAVKNAIDGCVVPGAGAVEVMAEALIKHKSVKRAQLGVQAFDALIIPKV
LAONGSPDLOETLVKIRTEHSESGQLVGVDLNTGSEPMVAEVDNVCVKQLHASC
TVIATNILLVDEIMRAGMSLKG"

ORIGIN

Query Match 1.6%; Score 21; DB 4; Length 1596;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTGTGCTG 1135
Db 1217 GTGTGTTCCAGTGTGCTG 1237

RESULT 29
LOCUS HUMHTR3A 1685 bp mRNA linear PRI 31-DEC-1994
DEFINITION Human chaperonin-like protein (HTR3) mRNA, complete cds.
ACCESSION M94083
VERSION M94083.1 GI:184461
KEYWORDS chaperonin-like protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Segel, G.B., Boal, T.R., Cardillo, T.S., Murant, F.G., Lichtman, M.A.
and Sherman, F.
Isolation of a gene encoding a chaperonin-like protein by
complementation of yeast amino acid transport mutants with human
cDNA
Proc. Natl. Acad. Sci. U.S.A. 89 (13), 6060-6064 (1992)
JOURNAL MEDLINE 92335237
COMMENT PUBMED 1352881
FEATURES
source Original source text: Homo sapiens cDNA to mRNA.
location/Qualifiers
1..1685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="b-lymphocyte"
1..1685
/gene="HTR3"
162..1331
/gene="HTR3"
/codon_start=1
/product="chaperonin-like protein"
/protein_id="AA58676.1"
/translation="MDRETLIDVARTSLRTKVHAEADVLTEAVVDSILAIKKODEPI
DLFMIEMEMKSEDTSLIRGLVDHGARRPMKRRVEDAYILTCNVSLSEYKTEV
NSGFPFKSAEERKVLKARERKTFEDVKKIVE LKKVCGSDGKGFVIVNKGIDPESDLAKGIVALKRARRNNERLTLACGVALN
SLDDINPDLGHAAGLYETLTSEKFTLEKCNPRSVTLVKGPKHTLTQIKAIR
DGLRAVKNAIDGCVVPGAGAVEVMAEALIKHKSVKRAQLGVQAFDALIIPKV
LAONGSPDLOETLVKIRTEHSESGQLVGVDLNTGSEPMVAEVDNVCVKQLHASC
TVIATNILLVDEIMRAGMSLKG"

ORIGIN

Query Match 1.6%; Score 21; DB 9; Length 1685;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTGTGCTG 1135
Db 952 GTGTGTTCCAGTGTGCTG 972

RESULT 30
LOCUS HUMTCP20 2010 bp mRNA linear PRI 13-JAN-1995
DEFINITION Human chaperonin protein (tcp20) gene complete cds.
ACCESSION U27706
VERSION U27706.1 GI:517064
KEYWORDS Chaperonin; TCP20; TRIC; chaperonin-like protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Segel, G.B., Boal, T.R., Cardillo, T.S., Murant, F.G., Lichtman, M.A.
and Sherman, F.
Isolation of a gene encoding a chaperonin-like protein by
complementation of yeast amino acid transport mutants with human
cDNA
Proc. Natl. Acad. Sci. U.S.A. 89 (13), 6060-6064 (1992)
JOURNAL MEDLINE 92335237
COMMENT PUBMED 1352881
FEATURES
source Original source text: Homo sapiens cDNA to mRNA.
location/Qualifiers
1..2010
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2010
/gene="TCP20"
87..1682
/gene="TCP20"
/codon_start=1
/product="chaperonin-like protein"
/protein_id="AA61061.1"
/translation="MAAVKTLNPKAEVAPAAQALAVNISANGLODVLRTNLGPKGTM
KMLVSGAGDILKLTGDNVILHMOIQHPASLIARVATQDDITGDTSSNVLIIGEL
LKOADLYISEGLHPIITEGFEAAKEKALQVEIQKVSREMDRETLIDVARTSLRTKV
HAEADVLTEAVVDSILAIKKODEPIDLPEVMEKHSKSEDTSLIRGLVDHGARR
PMKRRVEDAYILTCNVSLSEYKTEVNSGFPFKSAEERKVLKARERKTFEDVKKIVE
LKKVCGSDGKGFVIVNKGIDPESDLAKGIVALKRARRNNERLTLACGVALN
SLDDINPDLGHAAGLYETLTSEKFTLEKCNPRSVTLVKGPKHTLTQIKAIR
DGLRAVKNAIDGCVVPGAGAVEVMAEALIKHKSVKRAQLGVQAFDALIIPKV
LAONGSPDLOETLVKIRTEHSESGQLVGVDLNTGSEPMVAEVDNVCVKQLHASC
TVIATNILLVDEIMRAGMSLKG"

ORIGIN

Query Match 1.6%; Score 21; DB 9; Length 2010;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTGTGCTG 1135
Db 1303 GTGTGTTCCAGTGTGCTG 1323

RESULT 31
LOCUS AF385084 2562 bp mRNA linear PRI 12-JUN-2001
DEFINITION Homo sapiens heat shock protein mRNA, complete cds.

```

ACCESSION   AF385084
VERSION      AF385084.1  GI:14348899
KEYWORDS
SOURCE
ORGANISM     Homo sapiens (human)
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS      Lee, Y.-K. and Yoo, Y.-D.
TITLE        Homo sapiens chaperonin mRNA sequence
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 2562)
AUTHORS      Lee, Y.-K. and Yoo, Y.-D.
TITLE        Direct Submission
JOURNAL      Submitted (21-MAY-2001) Laboratory of Experimental Therapeutics,
              Korea Cancer Center Hospital, Nowon-gu, Seoul 139-706, Korea
FEATURES
Source
1..2562
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /issue_type="head and neck cancer"
   /note="chaperonin"
   /codon_start=1
   /product="heat shock protein"
   /protein_id="AAK61354.1"
   /db_xref="GI:14348890"
   /translation="MAAVKTLNPKAVARAQAALAVNISARGLODVLRTNLGPKGTM
KMVSGAGDIKLTQDGNVLLHEMOIQHPASLIAYATADODITGDTSTNVLIGEL
LKQADLYISEGLHPRIITGEFPAKEKALQLEEVKVSREMBRETLIDVARTSLRTKV
HAELADVLTAAVVDSIIAIKKODEPIDLMEITMEKHSRDTSLIGLVLDHGARR
PKMKRVEDAYIILTCVNSLEYETKTEVNSGFYKSAERKLVKAERKTEBRVKKIIIE
LKRKCGDSDKGVVYNQKIDIPFSLDALSKGIVALKRKRNMERLTLAGCGVALN
SPDLSPLDCLGHAGLYEYETLGEKFTFIEKCNPRSVTLLIKGPKKTLTQIKAVR
DGLRAVKNALIDDCVPGAGAVEMAAELIKRKSIVKGRAGLGVQAFADALLIIPKV
LAQNSGFDETLVKIKQAEHSBSGQLVGVDLNTGEBPMVAAEVGMNDYCVKKQLLHSC
TVIATNIIILVDEIMRAGMSLKG"

ORIGIN
Query Match      1.6%; Score 21; DB 9; Length 2562;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1115 GTGTGGTTCAGGTGCTGCTG 1135
      |||||||
Db      1272 GTGTGTTCCAGTCTGCTG 1292

RESULT 32
AB063318      2647 bp mRNA linear PRI 09-AUG-2001
LOCUS         Homo sapiens MODP-2, MODP-3 mRNA for acute morphine dependence
DEFINITION    related protein 2, acute morphine dependence related peptide 3,
               complete cds.
ACCESSION     AB063318
VERSION       AB063318.1  GI:14517631
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS      Wang, H., Gao, X., Li, L., Wang, B., Huang, Y. and Han, J.
TITLE        Homo sapiens chaperonin (MODP) mRNA expressed in SH-SY5Y
              neuroblastoma cells
JOURNAL      Published Only in Database (2001)
REFERENCE    2 (bases 1 to 2647)
AUTHORS      Wang, H., Gao, X., Li, L., Wang, B., Huang, Y. and Han, J.
TITLE        Direct Submission
JOURNAL      Submitted (18-JUN-2001) Haoran Wang, Peking University,
              Neuroscience Research Institute, 38 Xueyuan Rd., Beijing 100083,
              China (E-mail: haorwang@263.net, URL: www.bjmu.edu.cn,

```

```

FEATURES
Source
Tel:86-10-62091150, Fax:86-10-82072207)
Location/Qualifiers
1..2647
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /chromosome="7"
   /map="p11.1"
   /clone="VD-7-1"
   /cell_line="SH-SY5Y"
   /cell_type="neuroblastoma"
   /country="China:Beijing"
   /note="This gene expression was upregulated in the acute
morphine dependence neurons, but not in the chronic
dependence neurons.-this gene was primarily cloned by using
DD-PCR"
137..1732
   /gene="MODP-2"
137..1732
   /gene="MODP-2"
   /function="unknown"
   /note="similar to Tric chaperonin reported in L27706 and
heat shock protein reported in AF385084"
   /codon_start=1
   /evidence=experimental
   /product="acute morphine dependence related protein 2"
   /protein_id="BAB61032.1"
   /db_xref="GI:14517632"
   /translation="MAAVKTLNPKAVARAQAALAVNISARGLODVLRTNLGPKGTM
KMVSGAGDIKLTQDGNVLLHEMOIQHPASLIAYATADODITGDTSTNVLIGEL
LKQADLYISEGLHPRIITGEFPAKEKALQLEEVKVSREMBRETLIDVARTSLRTKV
HAELADVLTAAVVDSIIAIKKODEPIDLMEITMEKHSRDTSLIGLVLDHGARR
PKMKRVEDAYIILTCVNSLEYETKTEVNSGFYKSAERKLVKAERKTEBRVKKIIIE
LKRKCGDSDKGVVYNQKIDIPFSLDALSKGIVALKRKRNMERLTLAGCGVALN
SPDLSPLDCLGHAGLYEYETLGEKFTFIEKCNPRSVTLLIKGPKKTLTQIKAVR
DGLRAVKNALIDDCVPGAGAVEMAAELIKRKSIVKGRAGLGVQAFADALLIIPKV
LAQNSGFDETLVKIKQAEHSBSGQLVGVDLNTGEBPMVAAEVGMNDYCVKKQLLHSC
TVIATNIIILVDEIMRAGMSLKG"
2357..2524
   /gene="MODP-3"
2357..2524
   /gene="MODP-3"
   /note="putative"
   /codon_start=1
   /evidence=not_experimental
   /product="acute morphine dependence related peptide 3"
   /protein_id="BAB61033.1"
   /db_xref="GI:14517633"
   /translation="MKEGVVTKGHGWKVFELVQLYFQASELRONIHKKVSVIIFS
PKLRFHEIIVYI"
2614..2619
   /note="putative"
   /evidence=not_experimental
2636
   /note="putative"
   /evidence=not_experimental

ORIGIN
Query Match      1.6%; Score 21; DB 9; Length 2647;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1115 GTGTGGTTCAGGTGCTGCTG 1135
      |||||||
Db      1353 GTGTGTTCCAGTCTGCTG 1373

RESULT 33
CO492936      2837 bp DNA linear PAT 30-JAN-2004
LOCUS         Sequence 24803 from Patent WO0160860.
DEFINITION    CO492936
ACCESSION     CO492936
VERSION       CO492936.1  GI:41458555

```


KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 24803 23-AUG-2001;
Milleium Predictive Medicine, Inc. (US)
FEATURES
source 1..2837
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 1.6%; Score 21; DB 6; Length 2837;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGCTTCAGCTGCTGCTG 1135
|||||
Db 1307 GTGTGCTTCAGCTGCTGCTG 1327

RESULT 34
AC020045/2 23130 bp DNA linear HTG 03-JAN-2000

LOCUS AC020045
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC020045
VERSION AC020045.1 GI:6664852
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 23130)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT This sequence was identified as CDM:10211847 by the submitter. For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source 1..23130
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN
Query Match 1.6%; Score 21; DB 2; Length 23130;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1259 GTGTACTTGAGAAAAACGAC 1279
|||||
Db 554 GTGTACTTGAGAAAAACGAC 534

RESULT 35
AC092101 38875 bp DNA linear PRI 31-JAN-2004
LOCUS AC092101
DEFINITION Homo sapiens BAC clone RP11-745J15 from 7, complete sequence.
ACCESSION AC092101
VERSION AC092101.5 GI:17647102
KEYWORDS HTG.

SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Hillier, L.W., Fulton, R.S., Fulton, L.A., Graves, T.A., Pepin, K.H., Wagner-McPherson, C., Layman, D., Maas, S.P., Jaeger, S., Walker, R., Wyie, K., Sekhon, M.C., Becker, M.C., O'Laughlin, M.D., Schaller, M.E., Powell, G.A., Delahunty, K.D., Miner, T.L., Nash, W.E., Cordes, M., Du, H., Sun, H., Edwards, J., Bradshaw-Cordum, H., All, J., Andrews, S., Isak, A., Vanbrunt, A., Nguyen, C., Du, F., Lamar, B., Courtney, L., Kalkbush, J., Ozersky, P., Bielicki, L., Scott, K., Holmes, A., Harting, R., Harris, A., Strong, C.M., Hou, S., Tomlinson, C., Dauphin-Kohlberg, S., Kozlowski-Reilly, A., Leonard, S., Rohlfing, T., Rock, S.M., Tin-Kolam, A.M., Abbott, A., Minx, P., Mardis, R., Strowmatt, C., Latreille, P., Miller, N., Johnson, D., Murray, J., Woessner, J.P., Wendt, M.C., Yang, S.P., Schultze, B.R., Wallis, J.W., Spieth, J., Bieri, T.A., Nelson, J.O., Berkowitz, N., Wohldmann, P.E., Cook, L.L., Hickenbotham, M.T., Eldred, J., Williams, D., Bedell, J.A., Mardis, E.R., Clifton, S.W., Chissole, S.L., Marra, M.A., Raymond, C., Haugen, E., Gillett, W., Zhou, Y., James, R., Phelps, K., Iadonoto, S., Bubb, K., Simms, E., Levy, R., Clendenning, J., Kaul, R., Kent, W.J., Furey, T.S., Baertsch, R.A., Brent, M.R., Keibler, E., Flitce, P., Bork, P., Suyama, M., Bailey, J.A., Portnoy, M.E., Torrents, D., Chinwalla, A.T., Gish, W.R., Eddy, S.R., McPherson, J.D., Olson, M.V., Eichler, E.E., Green, E.D., Waterston, R.H. and Wilson, R.K.
TITLE The DNA sequence of human chromosome 7
JOURNAL Nature 424 (6945), 157-164 (2003)
MEDLINE 22737999
PUBMED 12853948

REFERENCE
AUTHORS Abbott, A., Nguyen, C. and Spalding, L.
TITLE The sequence of Homo sapiens BAC clone RP11-745J15
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 38875)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 38875)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 38875)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 38875)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 7 (bases 1 to 38875)
AUTHORS Wilson, R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2004) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Dec 13, 2001 this sequence version replaced gi:16596665.

Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: saplen@wustl.edu

Summary Statistics

Center project name: H_NH0745J15

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, U.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-15K19 the clone sequenced to the right is RP4-725G10, 2000 bp overlap. Actual start of this clone is at base position 163356 of RP11-415M24 actual end is at base position 16502 of RP4-725G10.

Location/Qualifiers

FEATURES

source

```
1..38875
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="7"
  /map="7"
```

repeat_region

```
/clone="RP11-745J15"
/clone_1fb="RPCI-11"
20..313
  /rpt_family="Alu"
```

gene

```
414..5822
  /gene="CCT6A"
  join(414..524,783..942,1041..1123,1965..2061,2690..2837,
3228..3361,4168..4270,5104..5176,5434..5822)
```

mRNA

```
/gene="CCT6A"
  join(<414..524,783..942,1041..1123,1965..2061,2690..2837,
3228..3361,4168..4270,5104..5176,5434..5506)
```

CDS

```
/gene="CCT6A"
  /note="Homo sapiens chaperonin containing TCP1, subunit 6A
(zeta 1) (CCT6A), mRNA.; H.NH0745015.1
This gene was based on gi(4502642)
continues from H_NH0015K19.1"
```

```
/product="unknown"
```

```
/protein_id="AA07451.1"
```

```
/db_xref="GI:41472460"
```

```
/translation="LIRGLVIDHGRHPDMKRRVEDAYILTCNVLSYEKTEVNSGPF
```

```
YKSAERKLVAKRKPLIEDRVKKIIEIKRVCDSDSGFVNIQKIDFESLDLSK
```

```
EGIVARAKRRMRRLTLACGVALNSFDLSDCIAGLVYETLGEKPFIEIK
```

```
CNNPRSVTLIKGNKHTLTQIDAVSDGJRAVKNALDDGCVVGAAGAVNAEALI
```

```
KHKPSVGRAGLGVQAFADALLIIPKYLANSQSPDLOETIYKIQAEHSQGLVGDVL
```

```
NTGEPVAAEAGVDNVCVKQKLHSCVATNTILVDETMRAMSSLKQ"
```

repeat_region

```
/rpt_family="L1"
```

repeat_region

```
1426..1721
```

```
repeat_region      /rpt_family="Alu"
                    2110..2420
repeat_region      /rpt_family="Alu"
                    3795..4067
repeat_region      /rpt_family="Alu"
                    4072..4115
repeat_region      /rpt_family="(T)n"
                    5792..5919
                    /rpt_family="L1"
misc_feature        /note="CpG island (%GC=59.5, o/e=0.77, #CpGs=88)"
                    6746..23091
gene                /gene="DKFZP5611024"
                    join(6746..6799,10903..11059,15418..15532,16595..16639,
mRNA                17007..17157,19255..19310,20523..20607,21949..23091)
                    /gene="DKFZP5611024"
repeat_region      6987..7012
                    /rpt_family="L1"
repeat_region      7013..7331
                    /rpt_family="Alu"
repeat_region      7332..7640
                    /rpt_family="Alu"
repeat_region      7641..7790
                    /rpt_family="L1"
repeat_region      7821..7987
                    /rpt_family="L1"
repeat_region      8009..8031
                    /rpt_family="(TTTA)n"
repeat_region      8044..8176
                    /rpt_family="Alu"
repeat_region      8222..8533
                    /rpt_family="Alu"
repeat_region      8717..8856
                    /rpt_family="MIR"
repeat_region      8858..9032
                    /rpt_family="Alu"
repeat_region      9033..9340
                    /rpt_family="Alu"
repeat_region      9724..9858
                    /rpt_family="Alu"
repeat_region      9998..10031
                    /rpt_family="MIR"
repeat_region      10032..10341
                    /rpt_family="Alu"
repeat_region      10342..10417
                    /rpt_family="MIR"
repeat_region      10444..10755
                    /rpt_family="Alu"
                    join(10923..11059,15418..15532,16595..16639,17007..17157,
CDS                1..6?; Score 21; DB 9; Length 38875;
                    Best Local Similarity 100.0%; Pred. No. 31;
                    Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy                1115 GTGTGTTCCAGGTCTGCTG 1135
Db                3231 GTGTGTTCCAGGTCTGCTG 3251
RESULT 36
AC114896          115815 bp   DNA      linear   PLN 30-JUL-2004
LOCUS             AC114896
DEFINITION        Oryza sativa chromosome 3 BAC OSJNBa0091B22 genomic sequence.
ACCESSION         AC114896
VERSION           AC114896
KEYWORDS          complete sequence.
SOURCE            HTG.
ORGANISM          Oryza sativa (japonica cultivar-group)
                  Oryza sativa (japonica cultivar-group)
                  Bukarjota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                  Ehrhartoideae; Oryzae; Oryza.
REFERENCE         1 (baaes 1 to 115815)
```

AUTHORS Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Ganabinger, K., Jones, K.M., Overton, D.L., Telford, T., Kim, M.M., Bera, J.J., Jin, S.S., Fadon, D.W., Tallon, L.J., Koo, H., Ziemann, V., Hsiao, J., Blunt, S., Vanaken, S.S., Riedmuller, S.B., Uterbeck, T.T., Feldlyum, T.V., Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, J., White, O., Salzberg, S.L. and Fraser, C.M.

TITLE *Oryza sativa* chromosome 3 BAC OSJNBa0091B22 genomic sequence

JOURNAL Unpublished

AUTHORS 2 (bases 1 to 115815)

TITLE Buell, R.

JOURNAL Direct Submission

AUTHORS Submitted (13-MAR-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

TITLE 3 (bases 1 to 115815)

JOURNAL Buell, R.

AUTHORS Direct Submission

TITLE Submitted (16-OCT-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

JOURNAL 4 (bases 1 to 115815)

AUTHORS Buell, R.

TITLE Direct Submission

JOURNAL Submitted (30-JAN-2004) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

AUTHORS 5 (bases 1 to 115815)

TITLE Buell, R.

JOURNAL Direct Submission

AUTHORS Submitted (30-JUN-2004) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org

TITLE On Jan 30, 2004 this sequence version replaced gi:37693581.

JOURNAL Address all correspondence to: rice@tigr.org

COMMENT BAC clone OSJNBa0091B22 is from *Oryza sativa* chromosome 3. The orientation of the sequence is from SP6 to T7 end of the BAC clone. Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (<http://www.softberry.com/>), GenScan and GenScan+ (Chris Burge, <http://CCR-081.mlt.edu/GENSCAN.html>), GeneMarkHMM (Mark Borodovsky, <http://genemark.biology.gatech.edu/Genemark/>), and Geneslicer (Michael Pertea and Steven Salzberg, contact.mper@tigr.org), searches of the complete sequence against a peptide database and the plant EST database at TIGR (<http://www.tigr.org/tdb/cgi.shtml>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

This BAC overlaps with rice BAC OSJNBa0083F15 (AC133938) and OSJNBa0036E17 (AC099041).

FEATURES

source

1..115815

location/Qualifiers

organism="Oryza sativa (japonica cultivar-group)"

molecule="genomic DNA"

cultivar="Nipponbare"

db_xref="taxon:39947"

chromosome="3"

map="C10241S"

clone="OSJNBa0091B22"

note="japonica cultivar-group"

1044..1096

repeat_region

1044..1096

repeat_family="AT_rich"

4142..9190

gene

gene="OSJNBa0091B22.1"

note="contains Pfam profile: PF00075 RNase H"

join(4142..5250,5908..6046,6498..8246,8504..9190)

gene="OSJNBa0091B22.1"

CDS

join(4142..5250,5908..6046,6498..8246,8504..9190)

gene="OSJNBa0091B22.1"

/codon_start=1

/product="RNase H domain containing protein"

/protein_id="AAT81735.1"

/db_xref="gi:50836974"

/translation="MGFVSGIDDFVPPGOAFRFGSLDTFTNNFGKISLSDSDNSQSGRGQVAPFGIPSAIEPKISTELATNSNEIOYTPRPDDDDAAYPIILKLPDDLAAVFTTASSSPRRDPAPAPAIQSSSREVGITVPLGTGSTGRLQSSGEORRTEIVYDDPGRYDYHNDDPDEGEDVYTLRYFETIADNTEBQRQALREAEQRYRQEAERRLREERLQREQERRRRAAEADRQALREARRRRELGGQOYVEGRAVPRTPKQANAVATITLVTLKEDALQANRVNINIQTKTMAAASPVASARTPTGSRVPLRSQDYHQPSESVVAGSSSRSDHGRCSKPMQSRKSPSPHPSAPPLGSRRLRLQEAIVEPITRLASRKAIEBELKLAAGIKELALHDMANPLVWKTGQWICVDYVGLANKSPKDPGLRIQOVVYSTGCELSLDFDCSRVHOIRLKBDCIKTSPITPRFACVITMPFGKAKAGATYQRIQCFSTQIRNVEAYVDIVYKTKQKDLITDLEETPASTIRTFPMKLPKCKTIGVSGKLGVFVSHRGIQANPEKINAILNMPSPSQDVQALSTGCMALSRFVRLSERGMPFKLKTGNFQWGPQAQAFEDFKLLITPPVLASPHLOEPLILVYSATFOVSVTVLVEREEGHVQKQRPITVPSVLADSKTRYPQVQKLVGLVTLVRKLSHYFOQSHSTVTVLTFPGDILHNHENGIAKMALELMSDISFKPRTSISKQALADFVAEMTEFQDTPABKMEVYTMHFDGSKRLSTGAGVLTISPTEGRLSYVMIRFSASHVAEYFALRGKIAISLGRILRYGDSQVYVQVKNESCLDDNNTATRHVERKLEDFDGLLELTVLRHNNEADRNLANSKEEAPSDVFEHLVEPTVPRKEIVBATDQEARQGFEPVPLVSDADKIMRCEGQPLARQTHLPAQELQITPLSWPFAWGLDMIGPKKAIISGYTHLFWAIDKSKMIEAKPVATITADKAPFINIVLWFGVPRNRIITNNGTOFTGQFVDFDCEDFGIKICVASYAHPMNGQVERANGVTLQGIKAVFPRRLHPYAKWKEQLPSVLMSLRTTPSRATGQSSFFLYVGAEAMLPSREVPESLRFQNSFEHYG"

complement(9820..12736)

gene="OSJNBa0091B22.2"

/note="contains pfam profile: PF04195 putative gypay type transposon"

complement(join(49820..10163,10330..10443,10507..10696,10771..10886,11036..11394,11477..11744,12025..12736))

gene="OSJNBa0091B22.2"

complement(join(9820..10163,10330..10443,10507..10696,10771..10886,11036..11394,11477..11744,12025..12736))

gene="OSJNBa0091B22.2"

/codon_start=1

/product="hypothetical protein"

/protein_id="AAT81736.1"

/db_xref="gi:50836975"

/translation="MAERESFEGQMAPSDVTEBNLKEMVAGVLLAKETITGMRLACGEEFPPDHEVVVFSHPFYGSGSLPTSEFPGLIDFYGLSHLNPNSIVHTNFIHACAPLGVPRPALFGRITFLKPOPKMSQCVVAGGPGRLRGLTSGKYSMPKTSKNGWHAAPFYQNSALPESICLPVYQDTWNSIPKMDBAQALADYRMLKKEQSLQGEQKQIOTHFITKSRLAPTKERSHPPIAQRSPKQOTGASGPKIRSDAOPHTSOPTQDPSRKRKLIVSHDEDAKERTGSKGMQKQKQANPNKATASRPITPKIRSSSRPSDIDLSKDPEPTGTETGDSKERTGPAEDPSDDQATDVAN3SNKPTQNSADTVSDADQEPGQSGTEONKQDIPVEAQTNSPPTDAGDSNTPVRVQGGAPARPEITITVKMDDMOQDQIMINKQKQDEBVLKKTLSQATRLVNRHLRNKAKTATLEKVLPHGLTEATRDQLAEAKELAKTETHDRDRIAEQDENSEFBSGSKNTVQATISLEKRIKALEKHGELTKQRDSALKQVEAMFPNNPSSDASEITDKLCAIDTITFKNLKESMGGASIALAMTKSLYPKITIDIVADGFDGTSEBAALDLINDAQAAQKIAVVERFQDNYLPTPGFNSDDEKTEITD"

14484..14577

repeat_region

14484..14577

repeat_family="CT-rich"

20365..20420

repeat_region

20365..20420

repeat_family="CCG)n"

20487..20765

repeat_region

20487..20765

repeat_family="CCG)n"

21690..21721

repeat_region

21690..21721

repeat_family="AT_rich"

21870..21957

repeat_region

21870..21957

repeat_family="AT_rich"

22026..22053

repeat_region

22026..22053

repeat_family="AT_rich"

22254..22300

repeat_region

22254..22300

repeat_family="AT_rich"

22367..22413

repeat_region

22367..22413

repeat_family="AT_rich"

22710..22880

repeat_region

22710..22880

repeat_family="T-rich"

```

gene      complement(24488..26854)
          /gene="OSJNBa0091B22.3"
          /note="contains Pfam profile: PF00082 subtlase family"
mRNA
          complement(<24488..>26854)
          /gene="OSJNBa0091B22.3"
          complement(24488..26854)
          /gene="OSJNBa0091B22.3"
          /codon_start=1
          /product="subtlase family protein"
          /protein_id="AAR81739.1"
          /db_xref="GI:50838978"
          /translation="MSSFTLMINILVILALASLSPADALCYTHAGSGSVKREP
TPPPADAIRTYIVLVDPHPHGAATDDDHRRWHSFPLGGRMDGDQAIIISYT
EYFEGFARLUITAEIAGVHAKKGFYRAPRRRTLTMTHTPEFGLRGSPFMDV
AGYKGVVGLDITGVAAHPSPDDGVPPPARWRMGCAVAVATRCNNKIKVGSFV
DGGGGDDVDVGHGHTASTAGNFVAGSDRGLGAGTAGIAPGAHVAMVYVCSG
CDDAVLAGFDEAMKGVULSVSLGRMSSPPDEPRPIAASPAAKRTIVCAAG
GDBPRTVNDAPMLITVAAGVSGSFSTVILNGELVDGALAAQPSSTSYPLH
FSRKQPKNELAGIVDGVAGHIVVQSDPVEDSVSMATGAGVLIINTSEGY
TVLEBYGGMVQIVVAGGHNITEVARSSSSAGCKPNATVPFNTLLSVAPATVA
SSSRGSPKVAAGVLPDVLAQNLIAAMPPLHGGGGGGGLFKVTSRSMATPH
ASGVAAIVKSRHPDMLPAIKSAILITSDAVDAGNPILDEHHRATLFTAGHINP
ARAADPGIYDIADVADYICALLGDAGLGTIVNESISCGKLDKNIKPEQIANTPT
ITVPLPRSSSAAPPFTYNTKRTVTVNGPARRTYTAKLEIPRSLTKRVSPEKLVFSG
EKKGRSVTVSGGGGGGGEVSSLSWVGKHMSPVAVAPPYLKIGS"
repeat_region 24761..24800
          /rpt_family=" (CCG) n"
          /rpt_family=" (CCG) n"
repeat_region 25200..25223
          /rpt_family=" (CCG) n"
          /rpt_family=" (CCG) n"
repeat_region 26168..26198
          /rpt_family=" (CCG) n"
          /rpt_family=" (CCG) n"
repeat_region 28808..28839
          /rpt_family="GC rich"
          complement(29554..33847)
          /note="OSJNBa0091B22.4"
          /note="contains Pfam profile: PF00665 integrase core
domain"
mRNA      complement(join(<29554..30791,31005..31870,31974..32076,
32179..32500,32537..33554,33696..>33847))
          /gene="OSJNBa0091B22.4"
          complement(join(29554..30791,31005..31870,31974..32076,
32179..32500,32537..33554,33696..33847))
CDS
Query Match      1.6%; Score 21; DB 8; Length 115815;
Beet Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      702 GAAAAAATTGATTGAAAGTAA 722
          |||||||
Db      80309 GAAAAAATTGATTGAAAGTAA 80329
          |||||||

RESULT 37
AC012253/c      AC012253      128524 bp      DNA      linear      HTG 13-JUL-2000
LOCUS      Homo sapiens chromosome 11 clone RP11-19J13 map 11, LOW-PASS
DEFINITION      SEQUENCE SAMPLING.
ACCESSION      AC012253
VERSION      AC012253.3 GI:9123893
KEYWORDS      HTG; HTGS; PHASE0.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 128524)
AUTHORS      Bittren,B., Linton,L., Nussbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome 11, clone RP11-19J13
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 128524)
AUTHORS      Bittren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,U., Barma,N., Beckerly,R., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,

```

```

Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,U., Gaidyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lewoczky,J., Lien,C., Locke,K., Macdonald,P., Marguis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,U.,
Testaye,S., Tittel,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.U., Zimmer,A. and Zody,W.
Direct Submission
Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6715953.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L3734
Center clone name: 19_J_13
-----
* NOTE: This record contains 141 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1      843      942: contig of 842 bp in length
*      843      942: gap of 100 bp
*      943      1821: contig of 879 bp in length
*      1822      1921: gap of 100 bp
*      1922      2813: contig of 892 bp in length
*      2814      2913: gap of 100 bp
*      2914      3632: contig of 718 bp in length
*      3632      3731: gap of 100 bp
*      3732      4606: contig of 875 bp in length
*      4607      4706: gap of 100 bp
*      4707      5582: contig of 876 bp in length
*      5583      5682: gap of 100 bp
*      5683      6548: contig of 866 bp in length
*      6549      6648: gap of 100 bp
*      6648      7518: contig of 870 bp in length
*      7519      7618: gap of 100 bp
*      7619      8462: contig of 844 bp in length
*      8463      8562: gap of 100 bp
*      8563      9418: contig of 856 bp in length
*      9419      9518: gap of 100 bp
*      9519      10399: contig of 881 bp in length
*      10400      10499: gap of 100 bp
*      10500      11386: contig of 887 bp in length
*      11387      11486: gap of 100 bp
*      11487      12376: contig of 890 bp in length
*      12377      12476: gap of 100 bp
*      12477      13364: contig of 888 bp in length
*      13365      13464: gap of 100 bp
*      13465      14332: contig of 858 bp in length
*      14333      14422: gap of 100 bp
*      14423      15288: contig of 866 bp in length
*      15289      15388: gap of 100 bp
*      15389      16269: contig of 881 bp in length
*      16270      16369: gap of 100 bp
*      16370      17237: contig of 868 bp in length
*      17238      17337: gap of 100 bp

```

```

17338 18216: contig of 879 bp in length
18316 18316: gap of 100 bp
18317 19199: contig of 883 bp in length
19200 19299: gap of 100 bp
19300 20116: contig of 817 bp in length
20117 20216: gap of 100 bp
20217 21072: contig of 856 bp in length
21073 21172: gap of 100 bp
21173 22031: contig of 859 bp in length
22032 22131: gap of 100 bp
22132 23035: contig of 904 bp in length
23036 23135: gap of 100 bp
23136 23995: contig of 860 bp in length
23996 24095: gap of 100 bp
24096 24977: contig of 882 bp in length
24978 25077: gap of 100 bp
25078 25959: contig of 882 bp in length
25960 26059: gap of 100 bp
26060 26928: contig of 869 bp in length
26929 27028: gap of 100 bp
27029 27694: contig of 866 bp in length
27695 27994: gap of 100 bp
27995 28882: contig of 888 bp in length
28883 28982: gap of 100 bp
28983 29828: contig of 846 bp in length
29829 29928: gap of 100 bp
29929 30811: contig of 883 bp in length
30812 30911: gap of 100 bp
30912 31795: contig of 884 bp in length
31796 31895: gap of 100 bp
31896 32793: contig of 898 bp in length
32794 32893: gap of 100 bp
32894 33795: contig of 902 bp in length
33796 33895: gap of 100 bp
33896 34788: contig of 893 bp in length
34789 34888: gap of 100 bp
34889 35756: contig of 868 bp in length
35757 35856: gap of 100 bp
35857 36786: contig of 930 bp in length
36787 36886: gap of 100 bp
36887 37768: contig of 882 bp in length
37769 37868: gap of 100 bp
37869 38756: contig of 888 bp in length
38757 38856: gap of 100 bp
38857 39744: contig of 888 bp in length
39745 39844: gap of 100 bp
39845 40709: contig of 865 bp in length
40710 40809: gap of 100 bp
40810 41669: contig of 860 bp in length
41670 41769: gap of 100 bp
41770 42656: contig of 887 bp in length
42657 42756: gap of 100 bp
42757 43647: contig of 891 bp in length
43648 43747: gap of 100 bp
43748 44577: contig of 830 bp in length
44578 44677: gap of 100 bp
44679 45544: contig of 867 bp in length
45545 45644: gap of 100 bp
45645 46512: contig of 868 bp in length
46513 46612: gap of 100 bp
46613 47481: contig of 869 bp in length
47482 47581: gap of 100 bp
47582 48437: contig of 856 bp in length
48438 48537: gap of 100 bp
48539 49397: contig of 860 bp in length
49398 49497: gap of 100 bp
49498 50358: contig of 861 bp in length
50359 50458: gap of 100 bp
50459 51336: contig of 878 bp in length
51337 51436: gap of 100 bp
51437 52320: contig of 884 bp in length
52321 52420: gap of 100 bp
52421 53323: contig of 903 bp in length

```

```

* 53424 53423: gap of 100 bp
* 53424 53414: contig of 891 bp in length
* 54315 54414: gap of 100 bp
* 54415 55295: contig of 881 bp in length
* 55296 55395: gap of 100 bp
* 55395 56269: contig of 874 bp in length
* 56269 56369: gap of 100 bp
* 56369 57243: contig of 874 bp in length
* 57243 57344: gap of 100 bp
* 57344 58214: contig of 871 bp in length
* 58214 58314: gap of 100 bp
* 58315 59198: contig of 884 bp in length
* 59199 59298: gap of 100 bp
* 59299 60131: contig of 833 bp in length
* 60132 60232: gap of 100 bp
* 60232 61118: contig of 887 bp in length
* 61119 61218: gap of 100 bp
* 61219 62080: contig of 862 bp in length
* 62081 62180: gap of 100 bp
* 62181 63049: contig of 869 bp in length
* 63050 63149: gap of 100 bp
* 63150 64055: contig of 906 bp in length
* 64056 64155: gap of 100 bp
* 64156 65060: contig of 905 bp in length
* 65061 65160: gap of 100 bp
* 65161 66010: contig of 850 bp in length
* 66011 66110: gap of 100 bp
* 66111 66984: contig of 874 bp in length
* 66985 67084: gap of 100 bp
* 67085 67967: contig of 883 bp in length
* 67968 68067: gap of 100 bp
* 68068 68958: contig of 891 bp in length
* 69058 69058: gap of 100 bp
* 69059 69923: contig of 865 bp in length
* 69924 70023: gap of 100 bp

```

```

Query Match      1.6% Score 21; DB 2; Length 128524;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1260 TGTACTTGAGAGAAAAACGAGA 1280
DB      122485 TGTACTTGAGAGAAAAACGAGA 122465

RESULT 38
LOCUS   AL772361
DEFINITION Mouse DNA sequence from clone RP23-136C22 on chromosome 2, complete
sequence.
ACCESSION AL772361
VERSION   AL772361.5 GI:23380941
KEYWORDS  HTG.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 128685)
REFERENCE
  1 Tracey, A.
  Direct Submission
  Submitted (27-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  humquerry@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk
  On Sep 30, 2002 this sequence version replaced gi:22798220.
  ----- Genome Center
  Center: Wellcome Trust Sanger Institute
  Center code: SC
  Web site: http://www.sanger.ac.uk
  Contact: humquerry@sanger.ac.uk
  -----
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the

```

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPep; Information on the WormPep database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-136C22 is from the RPCT-23 Mouse PAC Library constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES
Source
Location/Qualifiers
1..128685
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-136C22"
/clone_1b="RPCT-23"

ORIGIN

Query Match 1.6%; Score 21; DB 10; Length 128685;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 768 GTTATTAAATGGGAGCTCA 788
Db 34054 GTTATTAAATGGGAGCTCA 34074

RESULT 39
AC096845/c
LOCUS AC096845 134105 bp DNA linear VRT 01-OCT-2002
DEFINITION Takifugu rubripes clone 214014, complete sequence.
AC096845
VERSION AC096845.2 GI:17386262
KEYWORDS HTG.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes

REFERENCE
AUTHORS Akheri N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B., Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S., Dietrich N.L., Granite S., Guan X., Gupta D., Haghighi P., Ho S.-L., Idol J.R., Karling E., Latic P., Lee-Lin S.-O., Legaspi R., Maduro Q.L., Maduro V.B., Masello C., Mastrian S.D., McCluskey J.C., McDowell J., Pearson R., Prasad A., Sherchenko Y., Statulicop S., Thomas J.W., Thomas P.J., Touchman J.W., Tsugeon C., Voet J.L., Walker M.A., Weethey K.D., Wiggins L., Zhang L.-H. and Green E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 134105)

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
SUBMITTED (28-SEP-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 134105)

REFERENCE
AUTHORS
TITLE
JOURNAL
SUBMITTED (06-DEC-2001) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
4 (bases 1 to 134105)
Green, E.D.
Direct Submission
Submitted (01-OCT-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
On Dec 6, 2001 this sequence version replaced gi:15799588.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@hgri.nih.gov
----- Project Information
Center project name: at1
Center clone name: 214014

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.

FEATURES
Source
Location/Qualifiers
1..134105
/organism="Takifugu rubripes"
/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="214014"
/clone_1b="Incyte Genomics"
115823..115825
/note="low quality single stranded/single chemistry
region"

misc_feature

QY 1120 GTTCCAGTGTCTGTGATAC 1140
Db 100041 GTTCCAGTGTCTGTGATAC 100021

Query Match 1.6%; Score 21; DB 5; Length 134105;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 40
AL732425/c
LOCUS AL732425 145042 bp DNA linear ROD 19-SEP-2002
DEFINITION Mouse DNA sequence from clone RP23-36411 on chromosome X, complete
sequence.
AL732425
VERSION AL732425.7 GI:23304648
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
SUBMITTED (18-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk clone requests: clonerequest@sanger.ac.uk
On Sep 23, 2002 this sequence version replaced gi:21711885.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

COMMENT
Submitted (06-DEC-2001) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-364L1 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6.

FEATURES

source

Location/Qualifiers
 1..145042
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="X"
 /clone="RP23-364L1"
 /clone_lib="RPCI-23"

ORIGIN

Query Match 1.6%; Score 21; DB 10; Length 145042;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1224 TGTCTACTGGGTGATTGGA 1244
 ||||||||||||||||||
 Db 25558 TGTCTACTGGGTGATTGGA 25538

Search completed: January 26, 2005, 09:09:50
 Job time : 3938 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 06:38:50 : Search time 484 Seconds
(without alignments)
14088.839 Million cell updates/sec

Title: US-10-069-062-6
Perfect score: 1299
Sequence: 1 atgcctaaagcattcagtcgc.....aagactataggtttatata 1299

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_23Sep04:*

1: geneseq1980a:***
2: geneseq1990a:***
3: geneseq2000a:***
4: geneseq2001a:***
5: geneseq2001b:***
6: geneseq2002a:***
7: geneseq2002b:***
8: geneseq2003a:***
9: geneseq2003b:***
10: geneseq2003c:***
11: geneseq2003d:***
12: geneseq2004a:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1299	100.0	1299	4	AAD02792
2	1299	100.0	1299	4	AAD02791
3	891	68.6	1299	4	AA523422
4	891	68.6	1299	6	AB231736
5	421	32.4	547	4	AAD02787
6	334	25.7	577	4	AAD02788
7	65	5.0	65	6	AB228940
8	65	5.0	90	4	AA523667
9	25	1.9	25	4	AAD02790
10	24	1.8	36	4	AAD02796
11	22	1.7	22	4	AA523789
12	22	1.7	22	6	AB230804
13	21	1.6	21	4	AAD02789
14	21	1.6	476	9	ACH30537
15	21	1.6	479	6	ABK45175
16	21	1.6	518	6	ABV86952
17	21	1.6	530	6	ABK45703
18	21	1.6	771	4	AB197097
19	21	1.6	1574	4	ABQ54714
20	21	1.6	1825	12	ADQ22869
21	21	1.6	2010	6	ABV78079

22	21	1.6	2461	10	ADC29939	Adc29939 Human nov
23	21	1.6	2562	12	AD182850	Ad182850 Human PRO
24	21	1.6	2562	12	AD019167	Ad019167 Human PRO
25	21	1.6	2562	12	AD060113	Ad060113 Human CCT
26	21	1.6	2562	12	AD060114	Ad060114 Human CCT
27	21	1.6	2594	3	AA15841	AA15841 Human pro
28	21	1.6	2647	12	AD060115	Ad060115 Human pro
29	21	1.6	2637	5	ABV24814	Abv24814 Human pro
30	21	1.6	16682	4	AAK79798	Aak79798 Human imm
31	21	1.6	16682	4	AAK70010	Aak70010 Human imm
32	21	1.6	21436	4	AAK79799	Aak79799 Human imm
33	21	1.6	21436	4	AAK70011	Aak70011 Human imm
34	21	1.5	33	4	AAD02795	Aad02795 Candida a
35	20	1.5	502	12	ACH70211	Ach70211 Human gen
36	20	1.5	541	4	AA181760	AA181760 Human pol
37	20	1.5	630	6	AB212786	Ab212786 Arabidops
38	20	1.5	46852	8	ABQ76676	Abq76676 Androgen
39	20	1.5	218336	8	ABQ76678	Abq76678 Androgen
40	19	1.5	355	6	AB188534	Ab188534 C maxmore
41	19	1.5	593	8	ACA04641	ACA04641 cDNA enco
42	19	1.5	995	6	ABQ15522	Abq15522 Oligonuc1
43	19	1.5	995	6	ABQ15522	Abq15522 Oligonuc1
44	19	1.5	1282	5	ABV23937	Abv23937 Human pro
45	19	1.5	1282	5	ABV29820	Abv29820 Human pro
46	19	1.5	4699	10	AD554957	Ad554957 Human gen
47	19	1.5	4699	10	AD554973	Ad554973 Human gen
48	19	1.5	4699	10	AD554965	Ad554965 Human gen
49	19	1.5	4699	10	AD559402	Ad559402 Human gen
50	19	1.5	4699	10	AD554953	Ad554953 Human gen
51	19	1.5	4699	10	AD561072	Ad561072 Human gen
52	19	1.5	4699	10	AD554961	Ad554961 Human gen
53	19	1.5	4699	10	AD554969	Ad554969 Human gen
54	19	1.5	5760	4	AA532472	AA532472 Human gen
55	19	1.5	5763	4	AA532471	AA532471 Human gen
56	19	1.5	6733	4	AB128838	Ab128838 Drosophi1
57	19	1.5	10769	12	ADP79308	Adp79308 Human muc
58	19	1.5	12181	9	AD561072	Ad561072 Drosophi1
59	19	1.5	96597	9	ADA02501	Ada02501 Mouse Bac
60	19	1.5	96597	10	AD872239	Ad872239 Mouse Bac
61	19	1.5	96597	10	AD895749	Ad895749 Mouse Bac
62	19	1.5	110000	6	ABA90521_01	ABA90521_01
63	18	1.4	107	8	AB279966	Ab279966 Potexv10c
64	18	1.4	109	2	AAV75985	AAv75985 Scaphyloc
65	18	1.4	182	8	AB279967	Ab279967 Potexv10c
66	18	1.4	266	4	AAH35148	AAh35148 Human col
67	18	1.4	315	4	AAK55949	AAk55949 Human imm
68	18	1.4	335	4	AA199252	AA199252 Human exc
69	18	1.4	335	5	ABA16938	ABa16938 Human nex
70	18	1.4	335	5	ABA16937	ABa16937 Human nex
71	18	1.4	335	5	AA163602	AA163602 Human kid
72	18	1.4	345	5	ABA12118	ABa12118 Human nex
73	18	1.4	351	2	AAV78389	AAv78389 Scaphyloc
74	18	1.4	400	2	AAV78244	AAv78244 Scaphyloc
75	18	1.4	400	2	AAV78244	AAv78244 Scaphyloc
76	18	1.4	439	10	AD032786	Ad032786 Human mit
77	18	1.4	455	6	AB179441	AB179441 Human ova
78	18	1.4	472	6	AB163885	AB163885 Breast ca
79	18	1.4	494	4	AA112422	AA112422 Probe #23
80	18	1.4	494	4	ABA54130	ABa54130 Human foe
81	18	1.4	494	4	AA133779	AA133779 Probe #24
82	18	1.4	494	4	AA102337	AA102337 Probe #23
83	18	1.4	494	4	ABA43674	ABa43674 Human bre
84	18	1.4	494	4	ABA23877	ABa23877 Probe #23
85	18	1.4	494	4	AAK27844	AAk27844 Human bon
86	18	1.4	494	4	AAK02400	AAk02400 Human liv
87	18	1.4	494	4	AB527425	AB527425 Human ltr
88	18	1.4	494	5	AA102337	AA102337 Probe #23
89	18	1.4	494	6	AB502298	AB502298 Human gen
90	18	1.4	503	8	ABV27965	ABv27965 Potexv10c
91	18	1.4	548	5	ABV57150	ABv57150 Human pro
92	18	1.4	616	5	AA174653	AA174653 Aspergill1
93	18	1.4	619	5	AA572296	AA572296 DNA encod
94	18	1.4	624	6	ABN69044	ABn69044 Streptoco

95	18	1.4	720	8	ACA43365	ACA43365	Prokaryot
96	18	1.4	767	10	ACF65623	ACF65623	Photobhab
97	18	1.4	768	8	ACF72709	ACF72709	Staphyloc
98	18	1.4	784	4	AAL24414	AAL24414	Human bre
99	18	1.4	1050	10	ADCO7909	ADCO7909	Rice DNA
100	18	1.4	1071	2	AAQ94349	AAQ94349	Firegment
101	18	1.4	1096	4	AAS02047	AAS02047	DNA encod
102	18	1.4	1106	4	AAH24224	AAH24224	Human oxi
103	18	1.4	1177	4	AAF27656	AAF27656	DNA encod
104	18	1.4	1189	3	AAC45416	AAC45416	Arabidops
105	18	1.4	1190	2	AAV24145	AAV24145	Homo sapi
106	18	1.4	1191	3	AAC39581	AAC39581	Arabidops
107	18	1.4	1235	6	ABQ54527	ABQ54527	Human ova
108	18	1.4	1431	10	ADF03706	ADF03706	Bacterial
109	18	1.4	1437	10	ADC66348	ADC66348	Human GPC
110	18	1.4	1653	3	AAC48014	AAC48014	Zea mays
111	18	1.4	2000	12	ADJ11350	ADJ11350	Plant CDN
112	18	1.4	2059	8	ABZ36145	ABZ36145	Human sec
113	18	1.4	2114	2	AAQ39178	AAQ39178	Truncated
114	18	1.4	2200	4	ABL19908	ABL19908	Drosophi
115	18	1.4	2200	2	AAQ39180	AAQ39180	Truncated
116	18	1.4	2229	3	AAC44981	AAC44981	Arabidops
117	18	1.4	3247	3	AAC79589	AAC79589	Virulence
118	18	1.4	3247	3	AACT9590	AACT9590	Virulence
119	18	1.4	3247	6	ABQ83466	ABQ83466	Pasteurel
120	18	1.4	3278	6	ABQ83467	ABQ83467	Pasteurel
121	18	1.4	3441	8	ACA36434	ACA36434	Human gen
122	18	1.4	3505	2	AAO06320	AAO06320	Sequence
123	18	1.4	3505	2	AAO39177	AAO39177	Full leng
124	18	1.4	4455	2	AAQ93177	AAQ93177	PVX ORF1
125	18	1.4	4953	4	AAI58361	AAI58361	Human pol
126	18	1.4	4953	5	AAS70662	AAS70662	DNA encod
127	18	1.4	4953	5	AAS91310	AAS91310	DNA encod
128	18	1.4	4953	5	ADQ98570	ADQ98570	DNA encod
129	18	1.4	4953	9	ADB48330	ADB48330	Novel hum
130	18	1.4	4953	2	AAV74454	AAV74454	Staphyloc
131	18	1.4	10320	2	AAK65037	AAK65037	Human imm
132	18	1.4	10320	2	AAK65037	AAK65037	Human imm
133	18	1.4	13630	4	AAS27709	AAS27709	DNA encod
134	18	1.4	13630	4	ABK44032	ABK44032	Genomic D
135	18	1.4	13630	10	ADB94512	ADB94512	Novel hum
136	18	1.4	19380	6	AAS61427	AAS61427	Human gen
137	18	1.4	19533	10	ADC87260	ADC87260	Human GPC
138	18	1.4	26110	4	AAK78526	AAK78526	Human imm
139	18	1.4	26110	4	AAK65036	AAK65036	Human imm
140	18	1.4	32195	4	AAK65036	AAK65036	Human car
141	18	1.4	32195	4	AAK65036	AAK65036	Human car
142	18	1.4	32195	4	AAK65036	AAK65036	Human DNA
143	18	1.4	32195	4	AAK65036	AAK65036	Human DNA
144	18	1.4	32195	6	ABQ66856	ABQ66856	Genomic D
145	18	1.4	32195	6	ABQ66856	ABQ66856	Human pol
146	18	1.4	32195	10	ADCI1143	ADCI1143	Human DNA
147	18	1.4	32195	10	ADCI1149	ADCI1149	Human DNA
148	18	1.4	32195	10	ADCI1149	ADCI1149	Human DNA
149	18	1.4	32195	12	ADIS5030	ADIS5030	Novel hum
150	18	1.4	32195	12	ADIS5030	ADIS5030	Novel hum
151	18	1.4	67674	12	ADP81772	ADP81772	Human imm
152	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
153	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
154	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
155	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
156	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
157	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
158	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
159	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
160	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
161	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
162	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
163	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
164	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
165	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
166	18	1.4	93273	9	AAI57580	AAI57580	Human MD-
167	18	1.4	93273	9	AAI57580	AAI57580	Human MD-

241	17	1.3	785	4	AAH08357	Aah08357 Human cDN*	314	17	1.3	1821	4	AA159865	AA159865 Human pol
242	17	1.3	832	6	ABO75358	Abq75358 Human lun	315	17	1.3	1830	4	AAH27068	Aah27068 Feline 1e
243	17	1.3	837	10	ADH84011	Adh84011 Enterococc	316	17	1.3	1833	4	AAH27069	Aah27069 Feline 1e
244	17	1.3	839	4	AA123909	AA123909 Human bre	317	17	1.3	1878	1	AAH70079	Aan70079 Cat recom
245	17	1.3	846	4	AAH06430	Aah06430 Human cDN	318	17	1.3	1878	2	AAO68884	Aag68884 Fely gp90
246	17	1.3	849	6	ABN66217	Abn66217 Streptoco	319	17	1.3	1929	12	AD003715	Ado03715 Feline 1e
247	17	1.3	882	10	ACF70675	Acf70675 Photorhab	320	17	1.3	1950	11	ADM36091	Adm36091 Bovine me
248	17	1.3	936	8	ACA20445	Aca20445 Prokaryot	321	17	1.3	1964	6	ABM50381	Abm50381 Human gen
249	17	1.3	936	10	ABX07088	Abx07088 S. pneumo	322	17	1.3	1979	5	AAO08560	Aao08560 Feline 1e
250	17	1.3	939	3	AAZ91869	Aaz91869 Streptoco	323	17	1.3	1989	2	AAV49261	Aav49261 Fely-B en
251	17	1.3	950	4	AAK61400	Aak61400 Human imm	324	17	1.3	1991	2	PAC34660	Pac34660 Arabidops
252	17	1.3	963	4	AA112819	AA112819 Probe #27	325	17	1.3	1993	3	AAO50614	Aao50614 Arabidops
253	17	1.3	963	4	ABAA5420	Abaa5420 Human foe	326	17	1.3	2000	8	ADAV73276	Ada73276 Rice gene
254	17	1.3	963	4	AA134174	AA134174 Probe #28	327	17	1.3	2000	12	ADJ41385	Adj41385 Plant cre
255	17	1.3	963	4	ABAA4065	Abaa4065 Human bre	328	17	1.3	2049	12	ADN73692	Adn73692 Thale cre
256	17	1.3	963	4	ABAA4303	Abaa4303 Probe #27	329	17	1.3	2070	2	AAO68881	Aao68881 Fely gp70
257	17	1.3	963	4	AAK28254	Aak28254 Human bon	330	17	1.3	2076	1	AAH70076	Aan70076 Cat recom
258	17	1.3	963	4	AAK02811	Aak02811 Human bra	331	17	1.3	2076	6	ABN68118	Abn68118 Streptoco
259	17	1.3	963	4	ABSA27857	Abss27857 Human liv	332	17	1.3	2079	8	ACA50324	Aca50324 Prokaryot
260	17	1.3	963	5	AA102736	AA102736 Probe #27	333	17	1.3	2086	5	AAAS6143	Aaas6143 DNA encod
261	17	1.3	963	6	ABSA02767	Abss02767 Human gen	334	17	1.3	2163	6	ABN67518	Abn67518 Streptoco
262	17	1.3	978	4	AA121998	AA121998 Probe #11	335	17	1.3	2166	8	ACA50621	Aca50621 Prokaryot
263	17	1.3	978	4	ABAA67072	Abaa67072 Human foe	336	17	1.3	2179	10	AD140424	Ad140424 Human pur
264	17	1.3	978	4	AA147288	AA147288 Probe #15	337	17	1.3	2179	12	ADN05441	Adn05441 Antipsoxi
265	17	1.3	978	4	ABAA49159	Abaa49159 Human bre	338	17	1.3	2195	2	AAV52838	Aav52838 Streptoco
266	17	1.3	978	4	ABAA4165	Abaa4165 Probe #12	339	17	1.3	2195	2	AAV33973	Aav33973 S. pneumo
267	17	1.3	978	4	AAK41246	Aak41246 Human bon	340	17	1.3	2232	6	ABK35916	Abk35916 cDNA sequ
268	17	1.3	978	4	AAK15516	Aak15516 Human bra	341	17	1.3	2236	2	AAO50419	Aao50419 Partial B
269	17	1.3	978	4	ABSA40846	Abss40846 Human liv	342	17	1.3	2339	4	AAH52435	Aah52435 S. epider
270	17	1.3	978	5	AA107693	AA107693 Probe #76	343	17	1.3	2398	3	AAO50327	Aao50327 Arabidops
271	17	1.3	978	6	ABSA15257	Abss15257 Human gen	344	17	1.3	2400	3	AAO38222	Aao38222 Arabidops
272	17	1.3	998	6	ABT07198	ABT07198 Human Cpg	345	17	1.3	2415	9	ADA30132	Ada30132 DNA encod
273	17	1.3	1001	3	AAH51289	Aah51289 Human GLC	346	17	1.3	2468	6	ABA92931	Abag2931 Human pro
274	17	1.3	1001	3	AAH51286	Aah51286 Human GLC	347	17	1.3	2474	12	ADO35685	Ado35685 Novel mou
275	17	1.3	1001	3	AAH51286	Aah51286 Human GLC	348	17	1.3	2483	2	AA113266	Aat13266 Fely prov
276	17	1.3	1001	3	AAH51286	Aah51286 Human GLC	349	17	1.3	2484	2	AAO05253	Aao05253 Sequence
277	17	1.3	1020	8	ACA30308	Aca30308 Prokaryot	350	17	1.3	2499	2	AAQ02913	Aaq02913 Fely-B en
278	17	1.3	1028	10	ABE09710	ABE09710 Novel DNA	351	17	1.3	2523	2	AAH76140	Aah76140 Feline 1e
279	17	1.3	1039	8	ABE20406	ABE20406 Oncofaeta	352	17	1.3	2524	4	AAH33501	Aah33501 Human col
280	17	1.3	1104	2	AAK36181	Aak36181 Human TCA	353	17	1.3	2530	6	ABK87112	Abk87112 cDNA enco
281	17	1.3	1158	2	AAAC48332	Aaac48332 Arabidops	354	17	1.3	2616	4	AB130240	Ab130240 Drosophi
282	17	1.3	1162	2	AAH84813	Aah84813 Gentiana	355	17	1.3	2679	8	ACA46777	Aca46777 Prokaryot
283	17	1.3	1276	3	AACT9777	AACT9777 Human sec	356	17	1.3	2709	6	ABN28857	Abn28857 Staphyloc
284	17	1.3	1293	1	AAH70078	AAH70078 Cat recom	357	17	1.3	2844	3	AAZ50876	Aaz50876 Sheep GBS
285	17	1.3	1293	2	AAOC6883	AAOC6883 Fely gp70	358	17	1.3	2844	4	AAO10326	Aao10326 Sheep gro
286	17	1.3	1305	6	AAAD30139	AAAD30139 Human pit	359	17	1.3	2878	10	ADD48133	Add48133 Rat gene
287	17	1.3	1305	6	AAAD30963	AAAD30963 Human PTT	360	17	1.3	2904	10	ADF81889	Adf81889 Leukemia
288	17	1.3	1305	6	ABN87445	Abn87445 Human PTT	361	17	1.3	2971	4	AA102959	Aa102959 Human rep
289	17	1.3	1305	6	ABAA98079	Abaa98079 Human PTT	362	17	1.3	2971	4	AA102843	Aa102843 Human rep
290	17	1.3	1323	6	ABAA91278	Abaa91278 Cryptospor	363	17	1.3	2971	8	ADA41564	Ada41564 Human sec
291	17	1.3	1350	4	AAK68474	Aak68474 Human imm	364	17	1.3	2971	10	ADC74653	Adc74653 Human sec
292	17	1.3	1356	8	ABT19236	ABT19236 Aspergill	365	17	1.3	2971	10	ADD38121	Add38121 cDNA clon
293	17	1.3	1357	6	ABT11403	ABT11403 Yeast sel	366	17	1.3	2971	10	ADA547697	Ada547697 BAC fragm
294	17	1.3	1396	12	ADF08204	ADF08204 Human BSN	367	17	1.3	2995	4	AAH54039	Aah54039 S. epider
295	17	1.3	1416	1	AAH70077	AAH70077 Cat recom	368	17	1.3	3137	4	AAAS22936	Aas22936 DNA encod
296	17	1.3	1419	2	AAO68882	AAO68882 Fely gp70	369	17	1.3	3191	4	AB123897	Ab123897 Drosophi
297	17	1.3	1440	12	AD003725	AD003725 Feline 1e	370	17	1.3	3226	9	ACF57427	Acf57427 Mouse N-a
298	17	1.3	1455	8	ABT21056	ABT21056 Aspergill	371	17	1.3	3301	4	AA166063	Aa166063 Seccharom
299	17	1.3	1503	10	ACF68342	ACF68342 Photorhab	372	17	1.3	3301	12	ADMA47696	Adma47696 Polynucle
300	17	1.3	1509	8	ABT20458	ABT20458 Aspergill	373	17	1.3	3343	4	AAH14448	Aah14448 Human cDN
301	17	1.3	1509	8	ABT18642	ABT18642 Aspergill	374	17	1.3	3464	4	AAH54633	Aah54633 S. epider
302	17	1.3	1520	4	AAAS22830	AAAS22830 Human cDN	375	17	1.3	3478	10	ADB662703	Adb662703 Human cDN
303	17	1.3	1520	10	ADSO9483	ADSO9483 Novel DNA	376	17	1.3	3508	8	ABT18048	Abt18048 Aspergill
304	17	1.3	1520	10	ADSO9482	ADSO9482 Novel DNA	377	17	1.3	3509	8	ABT19862	Abt19862 Aspergill
305	17	1.3	1520	10	ADSO9481	ADSO9481 Novel DNA	378	17	1.3	3543	4	AAH54916	Aah54916 S. epider
306	17	1.3	1520	10	ADSO9480	ADSO9480 Novel DNA	379	17	1.3	3627	4	AAH54400	Aah54400 S. epider
307	17	1.3	1520	10	ADSO9479	ADSO9479 Novel DNA	380	17	1.3	3695	2	AAO32844	Aao32844 Rat micro
308	17	1.3	1527	12	ADSO9478	ADSO9478 Novel mou	381	17	1.3	3738	10	ADB79868	Adb79868 Rat micro
309	17	1.3	1590	6	AAAD37215	AAAD37215 Human 653	382	17	1.3	3747	2	AAO38758	Aao38758 Encodes B
310	17	1.3	1689	6	ABZ232352	Abz232352 Candida a	383	17	1.3	3762	2	AAO21074	Aao21074 Bovine SS
311	17	1.3	1692	6	AAZ92180	Aaz92180 N-acetylgl	384	17	1.3	3883	10	AAO21074	Aao21074 Bovine SS
312	17	1.3	1749	2	AAO50420	Aao50420 cdb gene	385	17	1.3	4122	6	AAO48283	Aao48283 Ehrlichia
313	17	1.3	1791	6	ABZ12573	Abz12573 Arabidops	386	17	1.3	4173	4	AB112075	Ab112075 Drosophi

387	17	1.3	4290	4	ABL20734	Ab120734 Drosophila	460	17	1.3	14886	8	ADA44263	Ada44263 Human sec
388	17	1.3	4312	4	ABL26644	Ab126644 Drosophila	461	17	1.3	14886	8	ADA44518	Ada44518 Human sec
C 389	17	1.3	4333	8	ABZ73964	Abz73964 Secreted	462	17	1.3	14886	10	ADC20948	Adc20948 Human sec
C 390	17	1.3	4333	10	ABZ67542	Abz67542 Human sec	463	17	1.3	14886	10	ABZ68052	Abz68052 Human sec
C 392	17	1.3	4624	10	ADL62328	Adl62328 Human cdn	C 464	17	1.3	15320	4	AAK78660	Aak78660 Human imm
C 393	17	1.3	4624	10	ADL16207	Adl16207 Human cdn	C 465	17	1.3	15320	4	AAK78662	Aak78662 Human imm
C 394	17	1.3	4924	4	AAH18656	Aah18656 Human cdn	C 466	17	1.3	17000	4	AAK90956	Aak90956 Human dig
C 395	17	1.3	5191	4	AAH12389	Aah12389 Drosophila	C 467	17	1.3	17000	4	AAK86254	Aak86254 Human dig
C 396	17	1.3	5265	5	AAH85296	Aah85296 DNA encod	C 468	17	1.3	17000	5	AAK31991	Aak31991 Human imm
C 397	17	1.3	5298	6	ABO67133	Abog67133 Human ang	C 469	17	1.3	17000	5	ABN90346	Abn90346 Human liv
C 398	17	1.3	5365	4	ABL30250	Ab130250 Drosophila	C 470	17	1.3	17000	11	ADJ15259	Adj15259 Human liv
C 399	17	1.3	5597	4	AAH98721	Aah98721 Human EST	C 471	17	1.3	17200	8	AAK37025	Aak37025 Human liv
C 400	17	1.3	5597	10	AD121403	Ad121403 Novel hum	C 472	17	1.3	17200	12	ADJ30763	Adj30763 Human mus
C 401	17	1.3	5597	10	AD121362	Ad121362 Novel hum	C 473	17	1.3	17373	6	ABK6435	Abk6435 Human mus
C 402	17	1.3	5599	4	AAK35950	Aak35950 Human car	C 474	17	1.3	18446	10	ADJ52575	Adj52575 Human mus
C 403	17	1.3	5599	10	ABAI15753	Abai15753 Human ner	C 475	17	1.3	18446	10	ADJ52575	Adj52575 Human mus
C 404	17	1.3	5611	10	ADP81966	Adp81966 Leukemia	C 476	17	1.3	19217	6	AAK6973	Aak6973 Caenorhab
C 405	17	1.3	5632	4	AAK23030	Aak23030 DNA encod	C 477	17	1.3	22813	4	AAK82016	Aak82016 Human B7-
C 406	17	1.3	5670	10	ADC30272	Adc30272 Human nov	C 478	17	1.3	22813	4	AAK86303	Aak86303 Human imm
C 407	17	1.3	5703	6	ABK36432	Abk36432 HIV Casse	C 479	17	1.3	22813	4	AAK65271	Aak65271 Human imm
C 408	17	1.3	5716	10	ADE71216	Ad71216 HIV Casse	C 480	17	1.3	22813	4	AAK87333	Aak87333 Human imm
C 409	17	1.3	5735	6	ABK40066	Abk40066 Human nov	C 481	17	1.3	31595	10	ADP81661	Adp81661 Human imm
C 410	17	1.3	5762	4	ABV20280	Abv20280 Human che	C 482	17	1.3	33718	4	AAK65257	Aak65257 Human imm
C 411	17	1.3	5762	4	ABV20280	Abv20280 Human pro	C 483	17	1.3	33718	4	AAK86411	Aak86411 Human imm
C 412	17	1.3	5997	5	ABV20280	Abv20280 Human pro	C 484	17	1.3	37590	4	AAK12439	Aak12439 Human imm
C 413	17	1.3	5999	5	ABV20280	Abv20280 Human pro	C 485	17	1.3	50335	9	AAK65890	Aak65890 Human imm
C 414	17	1.3	6005	5	ABX71306	Abx71306 Human uie	C 486	17	1.3	57728	9	AAK87588	Aak87588 Human tum
C 415	17	1.3	6106	4	AAK65270	Aak65270 Human uie	C 487	17	1.3	59065	6	ABD42416	Abd42416 Human gpi
C 416	17	1.3	6175	4	AAK65270	Aak65270 Human uie	C 488	17	1.3	59065	6	ABD41839	Abd41839 Human ser
C 417	17	1.3	6406	4	ABL26638	Ab126638 DNA encod	C 489	17	1.3	59065	10	ADG98545	Adg98545 Human liv
C 418	17	1.3	6406	4	ABL26638	Ab126638 DNA encod	C 490	17	1.3	76574	4	AAK28545	Aak28545 Human kln
C 419	17	1.3	6675	6	ABN80258	Abn80258 Human che	C 491	17	1.3	76574	4	ABL03438	Ab103438 Drosophila
C 420	17	1.3	6675	6	ABN80258	Abn80258 Human che	C 492	17	1.3	79590	11	ADJ27152	Adj27152 Human gen
C 421	17	1.3	6801	8	ABZ73858	Abz73858 Secreted	C 493	17	1.3	79590	11	ADJ27152	Adj27152 Human gen
C 422	17	1.3	6801	8	ABZ73858	Abz73858 Secreted	C 494	17	1.3	79684	9	ADA03074	Ada03074 Human hcg
C 423	17	1.3	6801	8	ABZ73858	Abz73858 Secreted	C 495	17	1.3	79684	9	ADA66358	Ada66358 Human hcg
C 424	17	1.3	7186	2	AAV52172	Aav52172 Streptococ	C 496	17	1.3	79684	10	ADB72812	Adb72812 Human hcg
C 425	17	1.3	7332	6	ABN72582	Abn72582 Streptococ	C 497	17	1.3	80374	12	ADM98956	Adm98956 Diterpene
C 426	17	1.3	7333	3	AAI58487	Aai58487 Human pol	C 498	17	1.3	81940	4	AAK05390	Aak05390 Human cit
C 427	17	1.3	7333	3	AAI58487	Aai58487 Human pol	C 499	17	1.3	81940	4	ABK64829	Abk64829 Human ben
C 428	17	1.3	7333	3	AAI58487	Aai58487 Human pol	C 500	17	1.3	81940	4	ADQ17315	Adq17315 Human scf
C 429	17	1.3	7653	5	ABA09696	Abao9696 Human bon	C 501	17	1.3	81940	4	ADQ17315	Adq17315 Human scf
C 430	17	1.3	7713	4	AB112074	Ab112074 Drosophila	C 502	17	1.3	81940	4	ADQ17315	Adq17315 Human scf
C 431	17	1.3	7713	4	AB112074	Ab112074 Drosophila	C 503	17	1.3	81940	4	ADQ17315	Adq17315 Human scf
C 432	17	1.3	7840	12	ADP67933	Adp67933 Human MDD	C 504	17	1.3	89378	12	ADN46845_20	Adn46845_20 Human WRN
C 433	17	1.3	8440	2	AAI13265	Aai13265 Fely F6a	C 505	17	1.3	89378	12	ADN46845_20	Adn46845_20 Human WRN
C 434	17	1.3	8440	2	AAI13265	Aai13265 Fely F6a	C 506	17	1.3	90583	8	ACD13447	Acd13447 Human DNA
C 435	17	1.3	9524	6	ABK39994	Abk39994 Human che	C 507	17	1.3	90798	12	ADP68859	Adp68859 Human chr
C 436	17	1.3	9524	6	ABK39994	Abk39994 Human che	C 508	17	1.3	91000	9	AAI61326	Aai61326 Human chr
C 437	17	1.3	9549	5	AAK63175	Aak63175 Human pur	C 509	17	1.3	93801	9	ABX13540	Abx13540 Human RGS
C 438	17	1.3	9567	6	ABK65017	Abk65017 Invertebr	C 510	17	1.3	94191	10	ADB11169_3	Adb11169_3 Human tlt
C 439	17	1.3	9567	6	ABK65017	Abk65017 Invertebr	C 511	17	1.3	96700	12	ADP08391	Adp08391 Human tlt
C 440	17	1.3	9829	6	AAI039237	Aai039237 Fruit fly	C 512	17	1.3	98686	12	ADP08391	Adp08391 Human tlt
C 441	17	1.3	10194	4	AB160324	Ab160324 Aphid let	C 513	17	1.3	98686	12	ADP08391	Adp08391 Human tlt
C 442	17	1.3	10194	4	AB160324	Ab160324 Aphid let	C 514	17	1.3	102644	10	ABK12169	Abk12169 Human tra
C 443	17	1.3	11071	10	ADC74657	Adc74657 Human sec	C 515	17	1.3	102644	10	ABK12169	Abk12169 Human tra
C 444	17	1.3	11071	10	ADC74657	Adc74657 Human sec	C 516	17	1.3	108316	10	ACF65378	Acf65378 Human DNA
C 445	17	1.3	11071	10	ADC74657	Adc74657 Human sec	C 517	17	1.3	110000	2	AAI42063_05	Aai42063_05 Human GPC
C 446	17	1.3	11079	5	AAK90954	Aak90954 Human dig	C 518	17	1.3	110000	6	ABN71527_01	Abn71527_01 Human chr
C 447	17	1.3	11079	5	AAK90954	Aak90954 Human dig	C 519	17	1.3	110000	6	ABN71527_01	Abn71527_01 Human chr
C 448	17	1.3	11079	5	AAK90954	Aak90954 Human dig	C 520	17	1.3	110000	6	ABN71527_01	Abn71527_01 Human chr
C 449	17	1.3	11427	6	ABN98897	Abn98897 Enterococ	C 521	17	1.3	110000	6	ABO74964_6	Abog74964_6 Human tlt
C 450	17	1.3	11427	6	ABN98897	Abn98897 Enterococ	C 522	17	1.3	110000	6	ABO74964_6	Abog74964_6 Human tlt
C 451	17	1.3	11456	4	AAK69400	Aak69400 Human imm	C 523	17	1.3	110000	6	ABO69245_05	Abog69245_05 Human tlt
C 452	17	1.3	12400	4	ABK07678	Abk07678 Drosophila	C 524	17	1.3	110000	6	ABO67196_0	Abog67196_0 Human tlt
C 453	17	1.3	14299	4	AAK90955	Aak90955 Human dig	C 525	17	1.3	110000	6	ABO69245_05	Abog69245_05 Human tlt
C 454	17	1.3	14299	5	AAK90955	Aak90955 Human dig	C 526	17	1.3	110000	6	ABO69245_05	Abog69245_05 Human tlt
C 455	17	1.3	14299	5	AAK90955	Aak90955 Human dig	C 527	17	1.3	110000	6	ABO69245_05	Abog69245_05 Human tlt
C 456	17	1.3	14299	11	ADJ15258	Adj15258 Human liv	C 528	17	1.3	110000	6	ABO69245_05	Abog69245_05 Human tlt
C 457	17	1.3	14886	8	ABZ73856	Abz73856 Secreted	C 529	17	1.3	110000	6	ABO67197_04	Abog67197_04 Human tlt
C 458	17	1.3	14886	8	ABZ73856	Abz73856 Secreted	C 530	17	1.3	110000	6	ABO67197_04	Abog67197_04 Human tlt
C 459	17	1.3	14886	8	ABZ73856	Abz73856 Secreted	C 531	17	1.3	110000	8	ABO53224_3	Abog53224_3 Human tlt
C 460	17	1.3	14886	8	ABZ73856	Abz73856 Secreted	C 532	17	1.3	110000	10	ABO53224_3	Abog53224_3 Human tlt

C 533	17	1.3	110000	10	ACF67367_10	Continuation (11 o	606	16	1.2	213	3	AA08431	AA08431 Peptostre
C 534	17	1.3	110000	10	ACF67367_11	Continuation (12 o	607	16	1.2	213	3	AA000384	AA000384 Human Muc
C 535	17	1.3	110000	10	ACF67367_17	Continuation (13 o	608	16	1.2	217	8	ACA13675	ACA13675 Prokaryot
C 536	17	1.3	110000	10	ACF65386_10	Continuation (11 o	609	16	1.2	219	2	AAV86960	AAV86960 EST clone
C 537	17	1.3	110000	10	ACF65384_4	Continuation (15 of	610	16	1.2	225	10	ADD49554	ADD49554 Human lun
C 538	17	1.3	110000	12	ADN25985_12	Continuation (13 o	611	16	1.2	227	12	ADJ41995	ADJ41995 Plant CDN
C 540	17	1.3	110000	12	ADN97989_12	Continuation (13 o	612	16	1.2	228	8	ABQ82891	ABQ82891 Human lun
C 540	17	1.3	110000	12	ADN47591_00	Continuation (13 o	613	16	1.2	231	2	AAI10683	AAI10683 Human lun
C 541	17	1.3	110000	12	ADN47209_00	Adna47209 Thermoco	614	16	1.2	242	9	ADA59104	ADA59104 Maize pho
C 542	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	615	16	1.2	242	9	ADA59104	ADA59104 Maize pho
C 543	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	616	16	1.2	247	3	AAI13210	AAI13210 Human bre
C 544	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	617	16	1.2	251	2	AAV90481	AAV90481 EST clone
C 545	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	618	16	1.2	254	5	AAH81249	AAH81249 Escherich
C 546	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	619	16	1.2	258	2	AAI10683	AAI10683 Human lun
C 547	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	620	16	1.2	258	2	AAI10683	AAI10683 Human lun
C 548	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	621	16	1.2	261	6	ABN16364	ABN16364 Human ORF
C 549	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	622	16	1.2	261	6	ABV86793	ABV86793 Human col
C 550	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	623	16	1.2	261	6	ABV86793	ABV86793 Human col
C 551	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	624	16	1.2	261	6	ABV86793	ABV86793 Human col
C 552	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	625	16	1.2	265	12	ACH85204	ACH85204 Human gen
C 553	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	626	16	1.2	269	3	AAH00125	AAH00125 Human col
C 554	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	627	16	1.2	272	6	ABR79583	ABR79583 Bacillus
C 555	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	628	16	1.2	275	12	ADP91979	ADP91979 Cotton ex
C 556	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	629	16	1.2	276	10	ADN99887	ADN99887 Bacterial
C 557	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	630	16	1.2	282	8	ACA14273	ACA14273 Prokaryot
C 558	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	631	16	1.2	288	8	ACF72857	ACF72857 Staphyloc
C 559	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	632	16	1.2	288	10	ADN85099	ADN85099 Farnesyl
C 560	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	633	16	1.2	291	2	AAI10685	AAI10685 Nucin-der
C 561	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	634	16	1.2	291	6	ABK80848	ABK80848 Bacillus
C 562	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	635	16	1.2	299	4	AAH73039	AAH73039 Human cer
C 563	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	636	16	1.2	299	4	AAH70329	AAH70329 Human cer
C 564	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	637	16	1.2	300	2	AAI213744	AAI213744 Human gen
C 565	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	638	16	1.2	300	5	AAH00390	AAH00390 Human col
C 566	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	639	16	1.2	300	5	AAH68755	AAH68755 Human lun
C 567	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	640	16	1.2	300	6	ABK38666	ABK38666 cDNA enco
C 568	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	641	16	1.2	300	8	ACA10995	ACA10995 Human lun
C 569	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	642	16	1.2	300	8	ABX99946	ABX99946 Lung canc
C 570	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	643	16	1.2	300	10	ADH46209	ADH46209 Human lun
C 571	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	644	16	1.2	300	12	ADH71961	ADH71961 Human lun
C 572	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	645	16	1.2	311	4	AAI21864	AAI21864 Probe #11
C 573	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	646	16	1.2	311	4	AAI47150	AAI47150 Probe #15
C 574	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	647	16	1.2	311	4	ABAA9017	ABAA9017 Human bre
C 575	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	648	16	1.2	311	4	ABAA4018	ABAA4018 Probe #12
C 576	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	649	16	1.2	311	4	AAK41101	AAK41101 Human bon
C 577	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	650	16	1.2	311	4	AAK15377	AAK15377 Human bra
C 578	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	651	16	1.2	311	4	ABSA40705	ABSA40705 Human liv
C 579	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	652	16	1.2	311	5	AAI07552	AAI07552 Probe #75
C 580	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	653	16	1.2	312	12	ADG99537	ADG99537 Kidney di
C 581	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	654	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 582	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	655	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 583	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	656	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 584	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	657	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 585	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	658	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 586	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	659	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 587	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	660	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 588	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	661	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 589	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	662	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 590	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	663	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 591	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	664	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 592	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	665	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 593	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	666	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 594	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	667	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 595	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	668	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 596	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	669	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 597	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	670	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 598	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	671	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 599	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	672	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 600	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	673	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 601	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	674	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 602	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	675	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 603	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	676	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 604	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	677	16	1.2	318	2	AAI06866	AAI06866 Nucin-der
C 605	17	1.3	110000	12	ADN47260_00	Adna47260 Thermoco	678	16	1.2	318	2	AAI06866	AAI06866 Nucin-der

C 679	16	1.2	339	4	AAK20190	Human bra	752	16	1.2	433	8	ABX47917
C 680	16	1.2	339	4	ABE45977	Human liv	753	16	1.2	434	5	ABV50179
C 681	16	1.2	339	6	ABE20569	Human gen	754	16	1.2	435	2	AAV52073
C 682	16	1.2	339	8	ACA47625	Prokaryot	755	16	1.2	435	3	AACT18047
C 683	16	1.2	342	10	ADH83121	Adh83121 Enterococ	756	16	1.2	443	6	ABE68867
C 684	16	1.2	345	6	ABG85898	Arabidops	757	16	1.2	444	2	AAK97731
C 685	16	1.2	346	2	AAV90298	EST clone	758	16	1.2	444	8	ABX37943
C 686	16	1.2	346	5	ABV50263	Human pro	759	16	1.2	444	9	ACH25653
C 687	16	1.2	350	12	ACH86626	Human gen	760	16	1.2	444	12	ADP18998
C 688	16	1.2	351	8	ACA43313	Prokaryot	761	16	1.2	445	4	ACH21943
C 689	16	1.2	357	8	AAE65197	Novel hum	762	16	1.2	448	8	ABX68318
C 690	16	1.2	358	5	AAE64718	Novel hum	763	16	1.2	449	8	ABX38899
C 691	16	1.2	360	9	ADK32638	DNA encod	764	16	1.2	455	9	ACH41726
C 692	16	1.2	361	12	ADM80016	Adm80016 Bovine BV	765	16	1.2	459	2	AAQ12772
C 693	16	1.2	369	4	AAI26308	Probe #16	766	16	1.2	461	12	ABX423378
C 694	16	1.2	369	4	AAI26308	Probe #16	767	16	1.2	462	8	ABE54061
C 695	16	1.2	369	4	AAI54533	Human foe	768	16	1.2	465	6	ABE933568
C 696	16	1.2	369	4	AAI54533	Probe #23	769	16	1.2	466	6	ABE964006
C 697	16	1.2	369	4	AAK48703	Human bon	770	16	1.2	468	4	AAK63510
C 698	16	1.2	369	4	AAK22536	Human bra	771	16	1.2	468	12	ACH85846
C 699	16	1.2	369	4	ABE48381	Human liv	772	16	1.2	470	9	ACH35057
C 700	16	1.2	369	6	ABE22427	Human gen	773	16	1.2	471	10	ADH85394
C 701	16	1.2	374	2	AAV90498	EST clone	774	16	1.2	471	10	ADH85394
C 702	16	1.2	378	3	AAE56731	Eucalyptu	775	16	1.2	472	4	AAE35597
C 703	16	1.2	387	10	ABZ40725	N. gonorr	776	16	1.2	472	10	ABE55676
C 704	16	1.2	390	2	AAV78341	Staphyloc	777	16	1.2	474	4	AAE28059
C 705	16	1.2	391	8	ABX54870	Bovine ES	778	16	1.2	474	9	ACH32732
C 706	16	1.2	400	2	AAV77940	Staphyloc	779	16	1.2	474	12	ADL04135
C 707	16	1.2	401	4	AAI81732	Human pol	780	16	1.2	476	6	ABE67297
C 708	16	1.2	402	3	AAE01530	Human sec	781	16	1.2	476	9	ACH27784
C 709	16	1.2	402	4	AAH57212	Human sto	782	16	1.2	476	10	ABE284721
C 710	16	1.2	402	9	ACH19541	Human adu	783	16	1.2	477	4	AAK61334
C 711	16	1.2	404	12	ADQ06672	Soybean t	784	16	1.2	477	10	ADBE9849
C 712	16	1.2	405	5	AAE65940	DNA encod	785	16	1.2	478	4	ABE44150
C 713	16	1.2	405	8	ABX46975	Bovine ES	786	16	1.2	478	4	ABE24384
C 714	16	1.2	405	8	ABX48409	Bovine ES	787	16	1.2	479	4	AAI14144
C 715	16	1.2	407	8	ABX43							

C 825	16	1.2	492	4	AB833064	AB833064 Human liv	898	16	1.2	550	12	ACH67577	ACH67577 Human gen
C 826	16	1.2	492	5	AA885198	AA885198 DNA encod	899	16	1.2	551	12	ACH76442	ACH76442 Human gen
C 827	16	1.2	492	6	AB808146	AB808146 Human gen	C 900	16	1.2	552	4	ACH71756	ACH71756 Human imm
C 828	16	1.2	493	6	ABN65639	ABN65639 Human can	C 901	16	1.2	552	9	ACH49376	ACH49376 Human lev
C 829	16	1.2	493	6	ABN50548	ABN50548 Human gen	C 902	16	1.2	552	9	AAH10327	AAH10327 Human can
C 830	16	1.2	496	4	AAH12760	AAH12760 Human CDN	C 903	16	1.2	553	4	AAH57038	AAH57038 Human muc
C 831	16	1.2	498	6	ABN67168	ABN67168 Streptococ	C 904	16	1.2	556	6	ABL36575	ABL36575 Human col
C 832	16	1.2	498	8	AB274130	AB274130 Secreted	C 905	16	1.2	557	2	AAQ24678	AAQ24678 H23-ETA-T
C 833	16	1.2	498	8	AB274131	AB274131 Secreted	C 906	16	1.2	558	2	AAQ24680	AAQ24680 H23-ETA-S
C 834	16	1.2	498	10	AB267688	AB267688 Human sec	C 907	16	1.2	559	4	AAI18607	AAI18607 Probe #85
C 835	16	1.2	498	10	AB267687	AB267687 Human sec	C 908	16	1.2	559	4	ABAB3606	ABAB3606 Human fce
C 836	16	1.2	502	3	AA656594	AA656594 Eucalyptu	C 909	16	1.2	559	4	AAI43722	AAI43722 Probe #12
C 837	16	1.2	509	3	AA652282	AA652282 Arabidops	C 910	16	1.2	559	4	ABA30804	ABA30804 Probe #92
C 838	16	1.2	510	6	ABG99600	ABG99600 Human cod	C 911	16	1.2	559	4	AAK37850	AAK37850 Human bon
C 839	16	1.2	513	3	AA656172	AA656172 Eucalyptu	C 912	16	1.2	559	4	AAK12134	AAK12134 Human bra
C 840	16	1.2	513	3	ABL27099	ABL27099 Drosophil	C 913	16	1.2	559	4	AB837479	AB837479 Human liv
C 841	16	1.2	514	5	ABV17458	ABV17458 Human pro	C 914	16	1.2	559	6	AB811843	AB811843 Human gen
C 842	16	1.2	518	8	ADA50583	ADA50583 Human muc	C 915	16	1.2	560	12	ACH70230	ACH70230 Human can
C 843	16	1.2	518	8	ADA57034	ADA57034 Human muc	C 916	16	1.2	563	6	ABN65787	ABN65787 Human can
C 844	16	1.2	519	5	ABX71323	ABX71323 Human sig	C 917	16	1.2	564	12	ACH70837	ACH70837 Human rec
C 845	16	1.2	520	5	ADL62477	ADL62477 Human ova	C 918	16	1.2	570	9	ADA09924	ADA09924 Human rec
C 846	16	1.2	522	6	ABN75836	ABN75836 Human den	C 919	16	1.2	571	4	AAI17131	AAI17131 Probe #70
C 847	16	1.2	524	6	ABZ16293	ABZ16293 Arabidops	C 920	16	1.2	571	4	ABAB1582	ABAB1582 Human fce
C 848	16	1.2	525	2	AAK27120	AAK27120 Calcineur	C 921	16	1.2	571	4	AAI41496	AAI41496 Probe #10
C 849	16	1.2	526	3	AA698386	AA698386 Human col	C 922	16	1.2	571	4	ABA29265	ABA29265 Probe #77
C 850	16	1.2	527	6	ABN62417	ABN62417 Human bre	C 923	16	1.2	571	4	AAK35779	AAK35779 Human bon
C 851	16	1.2	528	4	AAI14136	AAI14136 Human can	C 924	16	1.2	571	4	AAK09886	AAK09886 Human bra
C 852	16	1.2	529	4	ABL08923	ABL08923 Drosophil	C 925	16	1.2	571	4	AB835493	AB835493 Human liv
C 853	16	1.2	529	5	ABV57892	ABV57892 Human pro	C 926	16	1.2	571	6	AB810025	AB810025 Human gen
C 854	16	1.2	530	4	AA828950	AA828950 CDNA encod	C 927	16	1.2	571	12	ACH69491	ACH69491 Human gen
C 855	16	1.2	530	4	AA830187	AA830187 DNA encod	C 928	16	1.2	572	9	AA656951	AA656951 Human muc
C 856	16	1.2	530	4	AA834845	AA834845 CDNA encod	C 929	16	1.2	572	12	ACH72193	ACH72193 Human gen
C 857	16	1.2	530	4	AAI199584	AAI199584 Human exp	C 930	16	1.2	576	4	AA827488	AA827488 CDNA encod
C 858	16	1.2	530	4	AAI35644	AAI35644 Human mus	C 931	16	1.2	576	10	ADB93366	ADB93366 Human CDN
C 859	16	1.2	530	4	ABA06471	ABA06471 Human CDN	C 932	16	1.2	576	12	ACH71448	ACH71448 Human gen
C 860	16	1.2	530	4	ABK43875	ABK43875 DNA encod	C 933	16	1.2	577	5	AA875132	AA875132 DNA encod
C 861	16	1.2	530	5	AA829573	AA829573 Human end	C 934	16	1.2	577	8	AB273226	AB273226 Rice leaf
C 862	16	1.2	530	5	ADG15106	ADG15106 Human 7 t	C 935	16	1.2	579	3	AAAD00390	AAAD00390 Ubiquitin
C 863	16	1.2	530	6	ABV07785	ABV07785 Novel hum	C 936	16	1.2	580	5	ABV50121	ABV50121 Human pro
C 864	16	1.2	530	6	ABV83808	ABV83808 Human pol	C 937	16	1.2	585	4	AAI22998	AAI22998 Human bre
C 865	16	1.2	530	8	ACD01442	ACD01442 Human pol	C 938	16	1.2	586	6	ABN61329	ABN61329 Human can
C 866	16	1.2	530	8	ABX58632	ABX58632 CDNA encod	C 939	16	1.2	587	5	AD173368	AD173368 Human ova
C 867	16	1.2	530	10	ADCA6003	ADCA6003 Human neo	C 940	16	1.2	587	5	ADL38499	ADL38499 Human ova
C 868	16	1.2	530	12	AD154262	AD154262 CDNA encod	C 941	16	1.2	588	12	AD844412	AD844412 Human phe
C 869	16	1.2	530	12	ADJ28359	ADJ28359 Human mus	C 942	16	1.2	594	8	ACA332966	ACA332966 Prokaryot
C 870	16	1.2	531	16	ACA45429	ACA45429 Prokaryot	C 943	16	1.2	597	9	ADA29851	ADA29851 DNA encod
C 871	16	1.2	535	6	ABQ98850	ABQ98850 Human ORF	C 944	16	1.2	598	12	ACH72090	ACH72090 Human gen
C 872	16	1.2	535	8	ACC60422	ACC60422 Rice leaf	C 945	16	1.2	599	12	ACH72895	ACH72895 Human gen
C 873	16	1.2	535	10	ADD16915	ADD16915 DNA (Seq1	C 946	16	1.2	600	5	ABV57346	ABV57346 Human pro
C 874	16	1.2	535	10	ADK55990	ADK55990 Plant DNA	C 947	16	1.2	601	12	ADJ10890	ADJ10890 Recombina
C 875	16	1.2	536	5	ABV51261	ABV51261 Human pro	C 948	16	1.2	607	6	ABO57962	ABO57962 Human col
C 876	16	1.2	539	5	ABV51261	ABV51261 Human pro	C 949	16	1.2	609	6	ABN61486	ABN61486 Human can
C 877	16	1.2	539	8	ACA15224	ACA15224 Prokaryot	C 950	16	1.2	612	6	ABX73081	ABX73081 Human GPC
C 878	16	1.2	540	4	AAH35766	AAH35766 Human col	C 951	16	1.2	614	3	AAFO9012	AAFO9012 Fusarium
C 879	16	1.2	543	6	ABO92792	ABO92792 Triticum	C 952	16	1.2	615	5	ABV59167	ABV59167 Human pro
C 880	16	1.2	543	9	ADA50576	ADA50576 Mucin 1 (C 953	16	1.2	617	5	ADL41908	ADL41908 Human ova
C 881	16	1.2	543	9	AAAD56946	AAAD56946 Human muc	C 954	16	1.2	619	5	ADL44883	ADL44883 Human ova
C 882	16	1.2	546	5	AD176672	AD176672 Human ova	C 955	16	1.2	619	6	ABN65767	ABN65767 Human can
C 883	16	1.2	547	4	ABAB3244	ABAB3244 Human fce	C 956	16	1.2	623	3	AA637745	AA637745 Arabidops
C 884	16	1.2	547	4	AAI43318	AAI43318 Probe #12	C 957	16	1.2	624	3	AAZ890050	AAZ890050 Hydrophob
C 885	16	1.2	547	4	AAK37457	AAK37457 Human bon	C 958	16	1.2	626	8	ABZ53794	ABZ53794 Aspergill
C 886	16	1.2	547	4	AB837136	AB837136 Human liv	C 959	16	1.2	628	4	AAI19269	AAI19269 Human bre
C 887	16	1.2	548	10	ADDA7582	ADDA7582 Ret gene	C 960	16	1.2	634	5	ABAB17638	ABAB17638 Human ner
C 888	16	1.2	549	4	ABA60978	ABA60978 Human fce	C 961	16	1.2	634	6	ABN65812	ABN65812 Human can
C 889	16	1.2	549	4	AAI40875	AAI40875 Probe #95	C 962	16	1.2	635	5	ABAB17639	ABAB17639 Human ner
C 890	16	1.2	549	4	AAK80932	AAK80932 Human imm	C 963	16	1.2	636	5	AAK76487	AAK76487 Human imm
C 891	16	1.2	549	4	AAK80933	AAK80933 Human imm	C 964	16	1.2	638	5	ABV08410	ABV08410 Human CDN
C 892	16	1.2	549	4	ABA28923	ABA28923 Probe #73	C 965	16	1.2	638	4	AAI90192	AAI90192 Human pol
C 893	16	1.2	549	4	AAK35159	AAK35159 Human bon	C 966	16	1.2	640	3	AAAO1947	AAAO1947 Human col
C 894	16	1.2	549	4	AAK09270	AAK09270 Human bra	C 967	16	1.2	642	4	AAK52146	AAK52146 Human pol
C 895	16	1.2	549	4	AB834909	AB834909 Human liv	C 968	16	1.2	642	10	ADH83094	ADH83094 Bacterioc
C 896	16	1.2	549	6	AB809601	AB809601 Human gen	C 969	16	1.2	643	10	ADK60907	ADK60907 Ovarian c
C 897	16	1.2	549	6	AB809601	AB809601 Human gen	C 970	16	1.2	643	10	ADK60907	ADK60907 Ovarian c

C 971	16	1.2	645	4	ABK42395	Abk42395 Genomic s
C 972	16	1.2	645	4	ABK42396	Abk42396 Genomic s
C 973	16	1.2	645	9	ADB60552	ADB60552 Connectiv
C 974	16	1.2	645	9	ADB60551	ADB60551 Connectiv
C 975	16	1.2	646	12	ADN13004	Adn13004 Human pro
C 976	16	1.2	655	6	AA894961	AA894961 Human DNA
C 977	16	1.2	658	8	ACA22643	ACA22643 Prokaryot
C 978	16	1.2	660	10	ACA22643	ACA22643 Prokaryot
C 979	16	1.2	663	8	ACA29262	ACA29262 Prokaryot
C 980	16	1.2	665	3	AAFI4172	AAfi4172 Aspergill
C 981	16	1.2	673	5	AAH64776	AAh64776 Human sec
C 982	16	1.2	676	5	ADL63580	Adl63580 Human ova
C 983	16	1.2	677	4	ABLI0785	Abli0785 Drosophill
C 984	16	1.2	677	5	AD172909	Ad172909 Human ova
C 985	16	1.2	677	5	AD172909	Ad172909 Human ova
C 986	16	1.2	677	6	AAI37827	AAi37827 Soybean K
C 987	16	1.2	677	6	AAI37827	AAi37827 Soybean K
C 988	16	1.2	681	10	ADCT7295	Adct7295 DNA homol
C 989	16	1.2	681	10	ADCT6883	Adct6883 DNA homol
C 990	16	1.2	681	10	ADK57182	Adk57182 Plant DNA
C 991	16	1.2	681	10	ADK59429	Adk59429 Plant DNA
C 992	16	1.2	685	8	ADA26438	Ada26438 Human sec
C 993	16	1.2	685	8	ADA39792	Ada39792 Human sec
C 994	16	1.2	685	10	ACC50431	Acc50431 Human sec
C 995	16	1.2	685	10	ADCT3490	Adct3490 Human sec
C 996	16	1.2	685	10	ADA5981	Ada5981 Gene enco
C 997	16	1.2	685	12	ADL71499	Adl71499 Novel hum
C 998	16	1.2	711	2	AAZ15346	AAz15346 Human gen
C 999	16	1.2	711	2	AAFI3451	AAfi3451 Aspergill
C 1000	16	1.2	712	12	ADJ3482	Adj3482 plant cDN
			714	5	AA866691	AA866691 DNA encod

ALIGNMENTS

RESULT 1
AAD02792
ID AAD02792 standard; DNA; 1299 BP.

AC AAD02792;

DT 31-MAY-2001 (first entry)

DE Candida albicans phosphomevalonate kinase (PMK) coding sequence.

KM Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;

XX infection; ds.

OS Candida albicans.

XX Key

FT CDS 1..1299 Location/Qualifiers
FT CDS 1..1299 /tag= a
FT CDS 1..1299 /product= "C. albicans phosphomevalonate kinase (PMK)"
FT CDS 1..1299 /transl_except= (pos:730..732, aa:Glu)

XX WO200114533-A2.

XX 01-MAR-2001.

XX 15-AUG-2000; 2000MO-GB003100.

XX 21-AUG-1999; 99GB-00019766.

XX (ASTR) ASTRAZENECA AB.

XX (ASTR) ASTRAZENECA UK LTD.

XX Rosamond JDC, Schmeil NF;

XX WPI; 2001-218441/22.

XX P-PSDB; AAY72679.

XX New polypeptides and polynucleotides (ERG8) from Candida albicans, useful

PT	in assays for identifying inhibitors of phosphomevalonate kinase activity
PT	and as reagents for diagnosing C. albicans infection.
XX	Claim 6; Page 26; 29pp; English.
XX	The present sequence is phosphomevalonate kinase (PMK; ERG8) coding
CC	sequence from Candida albicans. The ERG8 protein is useful in an assay
CC	for identifying compounds that inhibit phosphomevalonate kinase (PMK)
CC	activity. These inhibitors are useful as anti-fungal agents. The ERG8 DNA
CC	and protein are also useful as reagents for diagnosing C. albicans
XX	infection
XX	Sequence 1299 BP; 439 A; 220 C; 273 G; 367 T; 0 U; 0 Other;
SO	Best Local Similarity 100.0%; Score 1299; DB 4; Length 1299;
Query Match	100.0%; Score 1299; DB 4; Length 1299;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1299; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCCTGCTGGTGAATATTGGTCTT 60
DB	1 ATGTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCCTGCTGGTGAATATTGGTCTT 60
QY	61 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGTCATCAGAAATGCAATGATTAACA 120
DB	61 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGTCATCAGAAATGCAATGATTAACA 120
QY	62 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGTCATCAGAAATGCAATGATTAACA 120
DB	62 GAGCCAAATTTATGATGCTTATGTGACAGCATTTGTCATCAGAAATGCAATGATTAACA 120
QY	121 CCAAAAGAACGATTTTAAAGAAATCTGAATCAAAATTTCTGACCCCAATTTGCAAC 180
DB	121 CCAAAAGAACGATTTTAAAGAAATCTGAATCAAAATTTCTGACCCCAATTTGCAAC 180
QY	122 CCAAAAGAACGATTTTAAAGAAATCTGAATCAAAATTTCTGACCCCAATTTGCAAC 180
DB	122 CCAAAAGAACGATTTTAAAGAAATCTGAATCAAAATTTCTGACCCCAATTTGCAAC 180
QY	181 GGAGATGGGAATATCATCATATCAATCAATCAAGAGCCAGAGAGTTCAGTCCGC 240
DB	181 GGAGATGGGAATATCATCATATCAATCAATCAAGAGCCAGAGAGTTCAGTCCGC 240
QY	182 GGAGATGGGAATATCATCATATCAATCAATCAAGAGCCAGAGAGTTCAGTCCGC 240
DB	182 GGAGATGGGAATATCATCATATCAATCAATCAAGAGCCAGAGAGTTCAGTCCGC 240
QY	241 ATAAATCCATTTTAAAGGCACTATATTCATCGTTTATGCTTATTTCAACCGACGAA 300
DB	241 ATAAATCCATTTTAAAGGCACTATATTCATCGTTTATGCTTATTTCAACCGACGAA 300
QY	242 ATAAATCCATTTTAAAGGCACTATATTCATCGTTTATGCTTATTTCAACCGACGAA 300
DB	242 ATAAATCCATTTTAAAGGCACTATATTCATCGTTTATGCTTATTTCAACCGACGAA 300
QY	301 GCATTTGATCTTGAATCATCATTTTCTCAGACCTCGGATATCATTCACAAAGATACT 360
DB	301 GCATTTGATCTTGAATCATCATTTTCTCAGACCTCGGATATCATTCACAAAGATACT 360
QY	302 GCATTTGATCTTGAATCATCATTTTCTCAGACCTCGGATATCATTCACAAAGATACT 360
DB	302 GCATTTGATCTTGAATCATCATTTTCTCAGACCTCGGATATCATTCACAAAGATACT 360
QY	361 GAAACCAAGACATCTCGAATGAGAAAACATTTCTTACCATTCGTCGTCATTAAC 420
DB	361 GAAACCAAGACATCTCGAATGAGAAAACATTTCTTACCATTCGTCGTCATTAAC 420
QY	362 GAAACCAAGACATCTCGAATGAGAAAACATTTCTTACCATTCGTCGTCATTAAC 420
DB	362 GAAACCAAGACATCTCGAATGAGAAAACATTTCTTACCATTCGTCGTCATTAAC 420
QY	421 GAAATGGAAGAACCGGATTAAGTTAGTGCAGAGATTAAGTGTCAAGTTGTCACAACT 480
DB	421 GAAATGGAAGAACCGGATTAAGTTAGTGCAGAGATTAAGTGTCAAGTTGTCACAACT 480
QY	422 GAAATGGAAGAACCGGATTAAGTTAGTGCAGAGATTAAGTGTCAAGTTGTCACAACT 480
DB	422 GAAATGGAAGAACCGGATTAAGTTAGTGCAGAGATTAAGTGTCAAGTTGTCACAACT 480
QY	481 TTATTTATCCATTTTATCCCAATGTTATCAGATTAAGTAAAGATTTTGGCAACGTT 540
DB	481 TTATTTATCCATTTTATCCCAATGTTATCAGATTAAGTAAAGATTTTGGCAACGTT 540
QY	482 TTATTTATCCATTTTATCCCAATGTTATCAGATTAAGTAAAGATTTTGGCAACGTT 540
DB	482 TTATTTATCCATTTTATCCCAATGTTATCAGATTAAGTAAAGATTTTGGCAACGTT 540
QY	541 GCACAGATTCACATTTGTTATGCCCAAAAAGATGATGATCGGTTGATGTTGCAACT 600
DB	541 GCACAGATTCACATTTGTTATGCCCAAAAAGATGATGATCGGTTGATGTTGCAACT 600
QY	542 GCACAGATTCACATTTGTTATGCCCAAAAAGATGATGATCGGTTGATGTTGCAACT 600
DB	542 GCACAGATTCACATTTGTTATGCCCAAAAAGATGATGATCGGTTGATGTTGCAACT 600
QY	601 GCAATTTATGCTGATGATGATTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 660
DB	601 GCAATTTATGCTGATGATGATTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 660
QY	602 GCAATTTATGCTGATGATGATTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 660
DB	602 GCAATTTATGCTGATGATGATTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 660
QY	661 CAGGTTCTGAAAGTATCTGAGAGTTCCTGAGAGTTCCTGAGAGTTCCTGAGAGTTC 720
DB	661 CAGGTTCTGAAAGTATCTGAGAGTTCCTGAGAGTTCCTGAGAGTTCCTGAGAGTTC 720
QY	662 CAGGTTCTGAAAGTATCTGAGAGTTCCTGAGAGTTCCTGAGAGTTCCTGAGAGTTC 720
DB	662 CAGGTTCTGAAAGTATCTGAGAGTTCCTGAGAGTTCCTGAGAGTTCCTGAGAGTTC 720
QY	721 AACTGGGAATTCAAACATGAAGATGATTAACATTCGAAATCAAGTATTAATGGGT 780
DB	721 AACTGGGAATTCAAACATGAAGATGATTAACATTCGAAATCAAGTATTAATGGGT 780
QY	722 AACTGGGAATTCAAACATGAAGATGATTAACATTCGAAATCAAGTATTAATGGGT 780
DB	722 AACTGGGAATTCAAACATGAAGATGATTAACATTCGAAATCAAGTATTAATGGGT 780
QY	781 GACGTCAAGGTGGCTCGAAGAACCCCAATTTGGTATCAAGATTCCTCAATGGAAAAAG 840
DB	781 GACGTCAAGGTGGCTCGAAGAACCCCAATTTGGTATCAAGATTCCTCAATGGAAAAAG 840

Qy	841	GAAGAACGAGAAAGAAAGCTGTTGTTGATATGACACAGCTTAATPAGGCCAATTTACAGTT	900
Db	841	GAAGAACGAGAAAGAAAGCTGTTGTTGATATGACACAGCTTAATPAGGCCAATTTACAGTT	900
Qy	901	ATGAAGAAATTAGGGAAATCGTGAAAAATACGATCAGACCCAGACCTTATATTAA	960
Db	901	ATGAAGAAATTAGGGAAATCGTGAAAAATACGATCAGACCCAGACCTTATATTAA	960
Qy	961	GAGTTAGATTCATTTCTGTTGACCTTTGACTGTGGGATTAAGAACATCAGAAAAGGTTA	1020
Db	961	GAGTTAGATTCATTTCTGTTGACCTTTGACTGTGGGATTAAGAACATCAGAAAAGGTTA	1020
Qy	1021	CAAGCATTAACACAAAATCAGAGGTTCCAAATTGAACCTGATGTCAAACCAAGTGTG	1080
Db	1021	CAAGCATTAACACAAAATCAGAGGTTCCAAATTGAACCTGATGTCAAACCAAGTGTG	1080
Qy	1081	GACCGTTCTCAAGAGATCTCTGTTGTGTGGTGCTGTGTTCCAGTGCTGTGGATAC	1140
Db	1081	GACCGTTCTCAAGAGATCTCTGTTGTGTGGTGCTGTGTTCCAGTGCTGTGGATAC	1140
Qy	1141	GATGCAATAGCTGATTAAGTGTGGAAAAATCAAGGGGAAATTTTAACACAGAAAATCTTT	1200
Db	1141	GATGCAATAGCTGATTAAGTGTGGAAAAATCAAGGGGAAATTTTAACACAGAAAATCTTT	1200
Qy	1201	GAAGAAATCCAGATTAATTTTCATATATGTTACTGGGTTGATTTGGAAAGACAAACGAAAGGT	1260
Db	1201	GAAGAAATCCAGATTAATTTTCATATATGTTACTGGGTTGATTTGGAAAGACAAACGAAAGGT	1260
Qy	1261	GTACTTGAAAGAAAACCGAAGACATATAGCTTATAA	1299
Db	1261	GTACTTGAAAGAAAACCGAAGACATATAGCTTATAA	1299

```

RESULT 2
AADO2791
ID AADO2791 standard; DNA; 1763 BP.
XX
XX AADO2791;
XX
XX 31-MAY-2001 (first entry)
XX
XX Candida albicans phosphomevalonate kinase (PMK) gene (ERG8 gene) .
XX
XX Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;
XX infection; ds.
XX
XX Candida albicans.
XX
XX Key Location/Qualifiers
XX FH 89..1387
XX CDS /*tag= a
XX FT /product= "C. albicans phosphomevalonate kinase (PMK)"
XX FT /transl_except= {pos:818..820, aa:Glu}
XX
XX WO200114533-A2.
XX
XX 01-MAR-2001.
XX
XX 15-AUG-2000; 2000MO-GB003100.
XX
XX 21-AUG-1999; 99GB-00019766.
XX
XX (ASTR ) ASTRAZENeca AB.
XX (ASTR ) ASTRAZENeca UK LTD.
XX
XX Rosamond JDC, Schnell NF;
XX
XX MPI; 2001-218441/22.
XX P-PSDB; AA772679.
XX
XX New polypeptides and polynucleotides (ERG8) from Candida albicans, useful
XX in assays for identifying inhibitors of phosphomevalonate kinase activity

```

PT	and as reagents for diagnosing <i>C. albicans</i> infection.				
XX					
PS	Claim 6, Fig 1, 29pp; English.				
XX					
CC	The present sequence is phosphomevalonate kinase (PMK) gene (ERG8 gene)				
CC	from <i>Candida albicans</i> . The ERG8 protein is useful in an assay for				
CC	identifying compounds that inhibit phosphomevalonate kinase (PMK)				
CC	activity. These inhibitors are useful as anti-fungal agents. The ERG8 DNA				
CC	and protein are also useful as reagents for diagnosing <i>C. albicans</i>				
CC	infection				
XX					
SQ	Sequence 1763 BP; 582 A; 300 C; 349 G; 530 T; 0 U; 2 Other;				
Query Match	100.0 %;	Score 1299;	DB 4;	Length 1763;	
Best Local Similarity	100.0 %;	Pred. NO. 0;			
Matches 1299;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0	

Query Match		100.0%;	Score 1299;	DB 4;	Length 1763;	
Best Local Similarity		100.0%;	Pred. No. 0;			
Matches 1299;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGTCAAAAGCATTTTAACTGACACCTGGAAAAGCATTTTCTGTGGTGAATATTTGGTTCTT	60			
Db	89	ATGTCAAAAGCATTTTAACTGACACCTGGAAAAGCATTTTCTGTGGTGAATATTTGGTTCTT	148			
Qy	61	GAGCCAAATTATATGATGTCTTAATGTGACGACGATTTGCATCAGAAAGCATGAGTATTAACA	120			
Db	149	GAGCCAAATTATATGATGTCTTAATGTGACGACGATTTGCATCAGAAAGCATGAGTATTAACA	208			
Qy	121	CCAAAAGAACCAATTGAAAGAAATCTAGAAATCAAAATTTCTTACACCCCAATTTGGCAAC	180			
Db	209	CCAAAAGAACCAATTGAAAGAAATCTAGAAATCAAAATTTCTTACACCCCAATTTGGCAAC	268			
Qy	181	GGAGATGGGAATATTCACATATTCATCAAATACAGAAAGCCAGAGAAATTCATCAGC	240			
Db	269	GGAGAAAGGGAATATTCACATATTCATCAAATACAGAAAGCCAGAGAAATTCATCAGC	328			
Qy	241	ATTAATATCATTTTATAGAGGCAACATATTCATGTTTATAGCTTATATTCACACGACGAA	300			
Db	329	ATTAATATCATTTTATAGAGGCAACATATTCATGTTTATAGCTTATATTCACACGACGAA	388			
Qy	301	GCAATTTGATCTTGAATATCATATTTACTCAGACCCTGGATATCTTACAGAAAGATATCT	360			
Db	389	GCAATTTGATCTTGAATATCATATTTACTCAGACCCTGGATATCTTACAGAAAGATATCT	448			
Qy	361	GAAGCAAGACATCTCTCGAATGAGAGAAAACATTTCTTTACCATCTCCGCGCAATTACC	420			
Db	449	GAAGCAAGACATCTCTCGAATGAGAGAAAACATTTCTTTACCATCTCCGCGCAATTACC	508			
Qy	421	GAAGTGGAAAAGACCGGATTAGGTTTCATCGCGAGAGATTAGTGTCAAGTGTGTCACAAGT	480			
Db	509	GAAGTGGAAAAGACCGGATTAGGTTTCATCGCGAGAGATTAGTGTCAAGTGTGTCACAAGT	568			
Qy	481	TTATATATCCCATTTTATATCCCAATGTTTATCATGATCGAATTAAGATATTTTGGACACGTT	540			
Db	569	TTATATATCCCATTTTATATCCCAATGTTTATCATGATCGAATTAAGATATTTTGGACACGTT	628			
Qy	541	GCACAGATTGCACATGTTTATGCCCCAAAAGATAGATCTGGGTTTGAATGGTGCACAT	600			
Db	629	GCACAGATTGCACATGTTTATGCCCCAAAAGATAGATCTGGGTTTGAATGGTGCACAT	688			
Qy	601	GCAATTTATGATCTGATTTGTATATAGAAAGTTTCAGCCAGCTTGTGATTAATGACGCTGTT	660			
Db	689	GCAATTTATGATCTGATTTGTATATAGAAAGTTTCAGCCAGCTTGTGATTAATGACGCTGTT	748			
Qy	661	CAGATTCTAAGATGTATCTCTGAGAAATTTCCCAACAGTTGAAAATTAATGATGAAAGT	720			
Db	749	CAGATTCTAAGATGTATCTCTGAGAAATTTCCCAACAGTTGAAAATTAATGATGAAAGT	808			
Qy	721	AACCTGGAAATTCACAAACATGAAAGATGTACATTCACATACGGAATCAAGTTATTAATGGGT	780			
Db	809	AACCTGGAAATTCACAAACATGAAAGATGTACATTCACATACGGAATCAAGTTATTAATGGGT	868			
Qy	781	GAGGTCAAGGATGGCTCAGAAAACACCCAAATTTGGTATCAGAGTATCTTCAATGGAAAAG	840			
Db	869	GAGGTCAAGGATGGCTCAGAAAACACCCAAATTTGGTATCAGAGTATCTTCAATGGAAAAG	928			

```
Qy 841 GAAAGCCAGAGAAAGCTCTGTGTATGACAGGTTATATGACCAATTTACGTTT 900
Db 929 GAAAGCCAGAGAAAGCTCTGTGTATGACAGGTTATATGACCAATTTACGTTT 988
Qy 901 ATGAGGAATTTGAGGAAATGCGTGAATAATACGACTCAGACCCAGAGATTATATATA 960
Db 989 ATGAGGAATTTGAGGAAATGCGTGAATAATACGACTCAGACCCAGAGATTATATATA 1048
Qy 961 GAGTTAGATCATCTGTGTAGCTTGTGACCTTGTGCAATTAAGAAATCGAAGAGGTTA 1020
Db 1049 GAGTTAGATCATCTGTGTAGCTTGTGACCTTGTGCAATTAAGAAATCGAAGAGGTTA 1108
Qy 1021 CAGGCAATTAACAAAATCAGAGGTTCCAAATTAAGATGTCGCAACCCAGTTGTTG 1080
Db 1109 CAGGCAATTAACAAAATCAGAGGTTCCAAATTAAGATGTCGCAACCCAGTTGTTG 1168
Qy 1081 GACCGTTGTCAAGAGATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 1169 GACCGTTGTCAAGAGATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1228
Qy 1141 GATGCAATAGCTGTATTTAGTTGTGAAATCAAGTGGGAAATTTTAAAGAGAAATCTT 1200
Db 1229 GATGCAATAGCTGTATTTAGTTGTGAAATCAAGTGGGAAATTTTAAAGAGAAATCTT 1288
Qy 1201 GAAATTCAGATTTATTTTATATATGTTTACTGCGTTGTTTGAAGAGCAAAAGAGT 1260
Db 1289 GAAATTCAGATTTATTTTATATATGTTTACTGCGTTGTTTGAAGAGCAAAAGAGT 1348
Qy 1261 GTACTTGAAGAAAACAGAGACTATATAGTTTATATA 1299
Db 1349 GTACTTGAAGAAAACAGAGACTATATAGTTTATATA 1387
```

RESULT 3

```
AAS23422
ID AAS23422 standard; DNA; 1299 BP.
```

```
XX AAS23422;
```

```
DT 04-DEC-2001 (first entry)
```

```
DE Candida albicans essential gene CAYMR220W (ERG8).
```

```
KM Gene identification; essential gene; GRAC3; pathogenic fungus;
```

```
XX gene replacement and conditional expression; fungal infection; ds.
```

```
OS Candida albicans.
```

```
PN WO200160975-A2.
```

```
PD 23-AUG-2001.
```

```
PF 20-FEB-2001; 2001WO-US00551.
```

```
PR 18-FEB-2000; 2000US-0183534P.
```

```
PA (ELITR-) ELITRA PHARM INC.
```

```
P1 Roemer T, Jiang B, Boone C, Bussey H;
```

```
XX WPI; 2001-489080/53.
```

```
DR P-PSDB; AAU15094.
```

```
PT Identifying genes essential to fungal metabolisms and identifying
```

```
XX potential therapeutic agents that target these genes.
```

```
XX Claim 22; Page 162; 324pp; English.
```

```
CC The present invention relates to novel methods for constructing fungal  
CC strains useful for identification and validation of gene products as  
CC targets for therapeutic agents, for creating a collection of identified  
CC essential genes, and screening assays for the discovery of new drugs. The  
CC invention provides the GRAC3 (gene replacement and conditional
```

```
CC expression) method for the construction of mutant organisms referred to  
CC as GRAC3 strains of the organism. The invention can be applied to any  
CC organism, particularly a pathogenic fungus e.g. Candida albicans,  
CC Aspergillus fumigatus and Cryptococcus neoformans. The methods are useful  
CC to identify agents that may be used in the treatment of fungal  
CC infections. AAS23381-AAS23442 represent C. albicans essential genes  
XX
```

```
SQ Sequence 1299 BP; 439 A; 221 C; 271 G; 368 T; 0 U; 0 Other;
```

```
Query Match 68.6%; Score 891; DB 4; Length 1299;
```

```
Best Local Similarity 99.4%; Pred. No. 0;
```

```
Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy 1 ATGTCAAAAGCAATTTAGTCACTGAGAAAGCAATTTCTGCGTGAATTTAGTGTCTT 60
Db 1 ATGTCAAAAGCAATTTAGTCACTGAGAAAGCAATTTCTGCGTGAATTTAGTGTCTT 60
Qy 61 GAGCCAAATTTAGATGCTTATATGATGACAGATTTGTCATCAGAAATGATGAGTTATACA 120
Db 61 GAGCCAAATTTAGATGCTTATATGATGACAGATTTGTCATCAGAAATGATGAGTTATACA 120
Qy 121 CCAAAAAGAACCAATTTGAAAGAAATCTGAATCAAAATTTCTGACCCCAATTTGCAAC 180
Db 121 CCAAAAAGAACCAATTTGAAAGAAATCTGAATCAAAATTTCTGACCCCAATTTGCAAC 180
Qy 181 GGAAGATGGGAATATCAATATCATCAAAATACAGAGAACCCAGAGAAATTCAGTCAGC 240
Db 181 GGAAGATGGGAATATCAATATCATCAAAATACAGAGAACCCAGAGAAATTCAGTCAGC 240
Qy 241 ATAAATCAATTTTGAAGGCAACTATATTCATGCTTTAGTTATTTTCAACCGACCGAA 300
Db 241 ATAAATCAATTTTGAAGGCAACTATATTCATGCTTTAGTTATTTTCAACCGACCGAA 300
Qy 301 GCATTTGATCTTGAATATCATTTACTCAGACCCCTGATATCATTCACAAGAGATACT 360
Db 301 GCATTTGATCTTGAATATCATTTACTCAGACCCCTGATATCATTCACAAGAGATACT 360
Qy 361 GAAACCAAGATCTCTGGAATGAGAGAAAACAATTTTTCATCATCTCTGTCATTTACC 420
Db 361 GAAACCAAGATCTCTGGAATGAGAGAAAACAATTTTTCATCATCTCTGTCATTTACC 420
Qy 421 GAAGTGAAGAAAGACCGGATTAAGTTCAATCGGAGAAATTAAGTGTCAAGT 480
Db 421 GAAGTGAAGAAAGACCGGATTAAGTTCAATCGGAGAAATTAAGTGTCAAGT 480
Qy 481 TTATTAATCCATTTTATCCCAATGTTATATAGATCAAGATTAAGTGTCAAGT 540
Db 481 TTATTAATCCATTTTATCCCAATGTTATATAGATCAAGATTAAGTGTCAAGT 540
Qy 541 GCACAGATTCACATTTGTTATGCCCAAAAAGATAGATCTGCGTTGATGTTGCAACT 600
Db 541 GCACAGATTCACATTTGTTATGCCCAAAAAGATAGATCTGCGTTGATGTTGCAACT 600
Qy 601 GCAATTTATGCTGATTTATATAGAAATTTACGCCAGCTTTGATTAATATGCGTTT 660
Db 601 GCAATTTATGCTGATTTATATAGAAATTTACGCCAGCTTTGATTAATATGCGTTT 660
Qy 661 CAGGTTCTGAAGAAAGTATCTCTGAGAGTCCCAAGAGTTGAAATTTGATTAAGT 720
Db 661 CAGGTTCTGAAGAAAGTATCTCTGAGAGTCCCAAGAGTTGAAATTTGATTAAGT 720
Qy 721 AACTGGGAATTCACATGAAGATGATCATTACCATACGAAATCAAGTTATTAAGT 780
Db 721 AACTGGGAATTCACATGAAGATGATCATTACCATACGAAATCAAGTTATTAAGT 780
Qy 781 GACGTCAAGGTGAGCTCAGAAACACCAATTTGATATACAGATCTCAATGCAAAAG 840
Db 781 GACGTCAAGGTGAGCTCAGAAACACCAATTTGATATACAGATCTCAATGCAAAAG 840
Qy 841 GAAAGCCAGAGAAAGCTCTGTGTATGACAGGTTATATGACCAATTTACGTTT 900
Db 841 GAAAGCCAGAGAAAGCTCTGTGTATGACAGGTTATATGACCAATTTACGTTT 900
```

QY 901 ATGAGGAATTTGAGGAAATGCGTGAATAATACGACTCAGACCCAGACTTATATTAA 960
DB 901 ATGAGGAATTTGAGGAAATGCGTGAATAATACGACTCAGACCCAGACTTATATTAA 960
QY 961 GAGTTAATCATCTTGTGAGCCTTTGACCTGTTGCGATTAAACATCAGAAAAGGTTA 1020
DB 961 GAGTTAATCATCTTGTGAGCCTTTGACCTGTTGCGATTAAACATCAGAAAAGGTTA 1020
QY 1021 CAAGCATTAACAAAAATCAGAGGTTCCATTGAACCTGATGTCCAAACCCAGTTGTTG 1080
DB 1021 CAAGCATTAACAAAAATCAGAGGTTCCATTGAACCTGATGTCCAAACCCAGTTGTTG 1080
QY 1081 GACCGTTGTCAGAGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
DB 1081 GACCGTTGTCAGAGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
QY 1141 GATGCAATAGCTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
DB 1141 GATGCAATAGCTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
QY 1201 GAAATATCAGATTATTTTCAATTAATGTTTACTGGGTTGTTGTTGTTGTTGTTGTTG 1260
DB 1201 GAAATATCAGATTATTTTCAATTAATGTTTACTGGGTTGTTGTTGTTGTTGTTGTTG 1260
QY 1261 GTACTTGAGAAAAACCAAGAACTATATAGTTTATTA 1299
DB 1261 GTACTTGAGAAAAACCAAGAACTATATAGTTTATTA 1299

RESULT 4
AB231736
ID AB231736 standard; DNA; 1299 BP.

XX AB231736;

XX 30-JAN-2003 (first entry)

XX Candida albicans essential gene SEQ ID NO 6023.

XX Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;

XX signal transduction; DNA replication; cell division; growth;

XX proliferation; Candida albicans; fungicide; antifungal; gene; ss.

XX Candida albicans.

XX WO200253728-A2.

XX 11-JUL-2002.

XX 26-DEC-2001; 2001WO-US049486.

XX 29-DEC-2000; 2000US-0259128P.

XX 20-FEB-2001; 2001US-00792024.

XX 22-AUG-2001; 2001US-0314050P.

XX (ELIT-) ELITRA PHARM INC.

XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;

XX WPI; 2002-566694/60.

XX P-PSDB; ABP73186.

XX Constructing strains for identifying gene products as effective targets

XX for therapeutic intervention, by inactivating in the strain one allele of

XX a gene and placing other allele of the gene under conditional expression.

XX Claim 37; SEQ ID NO 6023; 167bp + Sequence Listing; English.

CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival or growth of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC transnational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of an
CC essential Candida albicans gene used in the method of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office

XX Sequence 1299 BP; 439 A; 221 C; 271 G; 368 T; 0 U; 0 Other;

XX Query Match 68.6%; Score 891; DB 6; Length 1299;

XX Best Local Similarity 99.4%; Pred. No. 0;

XX Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTATGTCACCTGGAAAAGCATTTCTGCTGCGATATTTGGTTCTT 60

DB 1 ATGTCAAAAGCATTATGTCACCTGGAAAAGCATTTCTGCTGCGATATTTGGTTCTT 60

QY 61 GAGCAATTTATGATGCTTATGTCAGCATTTGTCATCAGAAATGATGAGTTATTA 120

DB 61 GAGCAATTTATGATGCTTATGTCAGCATTTGTCATCAGAAATGATGAGTTATTA 120

QY 121 CCAAAAAGAACGATTGGAAGATCTAGATCAAAATTTCTTACCCCAATTTGCAAC 180

DB 121 CCAAAAAGAACGATTGGAAGATCTAGATCAAAATTTCTTACCCCAATTTGCAAC 180

QY 181 GAGCATGGGAATATCATCATATCATCAATACAGAGAGCCAGAGATTCAGTCAGC 240

DB 181 GAGCATGGGAATATCATCATATCATCAATACAGAGAGCCAGAGATTCAGTCAGC 240

QY 241 ATAAATCATTTTATAGGCGACATATATCATGTTTATGCTTATATCAACGAGCGAA 300

DB 241 ATAAATCATTTTATAGGCGACATATATCATGTTTATGCTTATATCAACGAGCGAA 300

QY 301 GCATTTGATCTTGAATTCATCATTTTCTCAGACCTGATATGATTCACAAGAGATAT 360

DB 301 GCATTTGATCTTGAATTCATCATTTTCTCAGACCTGATATGATTCACAAGAGATAT 360

QY 361 GAAACCAAGACATCTCGATGAGAGAAAAACATTTCTTATCAATTCCTGCTTATAC 420

DB 361 GAAACCAAGACATCTCGATGAGAGAAAAACATTTCTTATCAATTCCTGCTTATAC 420

QY 421 GAAGTGGAAAAAGACCGGATTAGGTTGATGCGAGAGATTAGTCAAGTTGTCACAAGT 480

DB 421 GAAGTGGAAAAAGACCGGATTAGGTTGATGCGAGAGATTAGTCAAGTTGTCACAAGT 480

QY 481 TTATTAATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGCAACGTT 540

DB 481 TTATTAATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGCAACGTT 540

QY 541 GCAACAGTTGACATTTGTTATGCCCCAAAAAAGATAGATCTGGGTTGATGTTGCACT 600

DB 541 GCAACAGTTGACATTTGTTATGCCCCAAAAAAGATAGATCTGGGTTGATGTTGCACT 600

QY 601 GCAATTTATGCTGATGTTATGAGATTTTCAAGCACTTGTATTAATGACGTTT 660

DB 601 GCAATTTATGCTGATGTTATGAGATTTTCAAGCACTTGTATTAATGACGTTT 660

QY 661 CAGGTTCTAGAAAAGTATCTGGAAGTTCCCAAGAGTTGAAAAATGATGAAAGT 720

DB 661 CAGGTTCTAGAAAAGTATCTGGAAGTTCCCAAGAGTTGAAAAATGATGAAAGT 720

[illegible]

PT New polypeptides and polynucleotides (ERG8) from *Candida albicans*, useful
PT in assays for identifying inhibitors of phosphomemalonalate kinase activity
PT and as reagents for diagnosing *C. albicans* infection.
XX
PS Claim 6, Page 24, 29pp, English.

Claim 6; Page 24; 29pp; English.

The patent discloses phosphomethylatolactate kinase (PMK; ERG8) protein and their corresponding DNAs from *Candida albicans*. The ERG8 protein is useful in assays for identifying compounds that inhibit phosphomethylatolactate kinase (PMK) activity. These inhibitors are useful as antifungal agents. The ERG8 DNA and protein are also useful as reagents for diagnosing *C. albicans* infection. The present sequence is a *Candida albicans* DNA which is homologous to the *Saccharomyces cerevisiae* ERG8 gene.

SQ Sequence 547 BP; 184 A; 81 C; 123 G; 159 T; 0 U; 0 Other;

Query Match	Score	DB 4	Length
32.4%	421		547

2000 Locally Optimally 99.8%; Fred. NO. 4.1e-194;
 Matches 471; Conservative 0; Mismatches 1; Indels 0.

828 CCAATGAAAAGGAAAGCCAGAGAAGAGCTCTGTTGTATGACCAGCTTAATAGTGC 887

Db
1 CCAATGGAAGGAAAGCCAGAGAAGCTCTGTTGTGTATGACCAGCTTAATAGTGC 60

888 CAATTACAGTTTATGAAGGAATTGAGGGAATGCGTGAATAATACGACTCAGACCCAGA 947

Db 61 CAATTACAGTTTATGAAGGAATTGAGGGAATGCGTGAATAATACGACTCAGACCCAGA 120

948 GACTTATATTAAAGAGTTAGATCATTCCTGTGAGCCTTTGACTGTTGCCATTAGAACAT 1007

121 GAC IAAATTAAGAGTAAAGATCAATCTGTGAGCCCTTGAAGCTGTTGCGATTAGAACAT 180

2000 CAGAGGTTCAAATGAACTGATGTCCA 1067

-----TTCACCTTATTCAGAGGTCCTCAATTCGATGCTCCA 240

[illegible][illegible][illegible]

OY
1188 GCAGAAAAC TCTTGA AATCCAGA TTATTTC ATATA TCATTG ACCCACC CTTTCCC GGG

Db 361 GCAGAACTCTTGAATCCAGATTATTTTCATATGTTTACTGGTTCATTTCCACCA 420

1248 GCAAACGAGGTGTACTTGAAGAAAAACGAGACTATATAGGTTATAA 1299

421 GCAAACAGAGGTGTA
CTGAAGAAACCAGA
GACTATATAGTTATA
A 472

RESISTANCE

AD02788/c
ID AD02788 standard, nva. 577 55

XX
AC AAD02788:

31-MAY-2001 (first entry)

S. cerevisiae ERG8 gene homologous DNA #2 from *Candida albicans*

Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis:

1000

XXXXXXXXXXXX

01-MAR-2001


```
XX DE Tetracycline promoter downstream PCR primer (Tet-Down) #42.
XX XX
XX KW Gene identification; essential gene; GRACE; pathogenic fungus;
XX KW gene replacement and conditional expression; fungal infection;
XX KW PCR primer; Tet-Down; tetracycline promoter; ss.
XX OS Escherichia coli.
XX OS Candida albicans.
XX PN WO200160975-A2.
XX XX
XX PD 23-AUG-2001.
XX XX
XX PF 20-FEB-2001; 2001WO-US005551.
XX XX
XX PR 18-FEB-2000; 2000US-0183534P.
XX XX
XX PA (ELIT-) ELITRA PHARM INC.
XX XX
XX PI Roemer T, Jiang B, Boone C, Buesey H;
XX XX WPI; 2001-489080/53.
XX DR
XX XX
XX PT Identifying genes essential to fungal metabolisms and identifying
XX PT potential therapeutic agents that target these genes.
XX PS Disclosure; Page 297; 324pp; English.
XX XX
XX CC The present invention relates to novel methods for constructing fungal
XX CC strains useful for identification and validation of gene products as
XX CC targets for therapeutic agents, for creating a collection of identified
XX CC essential genes, and screening assays for the discovery of new drugs. The
XX CC invention provides the GRACE (gene replacement and conditional
XX CC expression) method for the construction of mutant organisms referred to
XX CC as GRACE strains of the organism. The invention can be applied to any
XX CC organism, particularly a pathogenic fungus e.g. Candida albicans,
XX CC Aspergillus fumigatus and Cryptococcus neoformans. The methods are useful
XX CC to identify agents that may be used in the treatment of fungal
XX CC infections. AAS23626-AAS23686 represent tetracycline promoter downstream
XX CC PCR primers (Tet-Down) used in the methods of the present invention
XX SQ
XX Sequence 90 BP; 28 A; 19 C; 18 G; 25 T; 0 U; 0 Other;
XX
XX Query Match 5.0%; Score 65; DB 4; Length 90;
XX Best Local Similarity 100.0%; Pred. No. 5.9e-21;
XX Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ATGTCAAAAGCATTAGTGACCTGGAAGCATTCTTGCTGGTGATATTGGTTCTT 60
XX DB 65 ATGTCAAAAGCATTAGTGACCTGGAAGCATTCTTGCTGGTGATATTGGTTCTT 6
XX
XX QY 61 GAGCC 65
XX DB 5 GAGCC 1
XX
XX RESULT 9
XX AAD02790/c
XX ID AAD02790 standard; DNA; 25 BP.
XX XX
XX AC AAD02790;
XX XX
XX DT 31-MAY-2001 (first entry)
XX XX
XX DE Candida albicans ERG8 gene isolating PCR primer #2.
XX KW Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;
XX KW infection; PCR primer; ss.
XX OS Candida albicans.
XX XX
XX PN WO200114533-A2.
```

```
XX XX
XX PD 01-MAR-2001.
XX XX
XX PF 15-AUG-2000; 2000WO-GB003100.
XX XX
XX PR 21-AUG-1999; 99GB-00019766.
XX XX
XX PA (ASTR ) ASTRAZENECA AB.
XX PA (ASTR ) ASTRAZENECA UK LTD.
XX PI Rosamond JDC, Schnell NF;
XX XX WPI; 2001-218441/22.
XX DR
XX XX
XX PT New polypeptides and polynucleotides (ERG8) from Candida albicans, useful
XX PT in assays for identifying inhibitors of phosphomevalonate kinase activity
XX PT and as reagents for diagnosing C. albicans infection.
XX PS Example 1; Page 25; 29pp; English.
XX XX
XX CC The patent discloses phosphomevalonate kinase (PMK; ERG8) protein and
XX CC their corresponding DNAs from Candida albicans. The ERG8 protein is
XX CC useful in assays for identifying compounds that inhibit phospho-
XX CC mevalonate kinase (PMK) activity. These inhibitors are useful as anti-
XX CC fungal agents. The ERG8 DNA and protein are also useful as reagents for
XX CC diagnosing C. albicans infection. The present DNA sequence is a PCR
XX CC primer which is used in the isolation of a clone containing the Candida
XX CC albicans ERG8 gene
XX SQ
XX Sequence 25 BP; 11 A; 1 C; 7 G; 6 T; 0 U; 0 Other;
XX
XX Query Match 1.9%; Score 25; DB 4; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 0.18;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 477 AAGTTATTATCCATTATTATCCCC 501
XX DB 25 AAGTTATTATCCATTATTATCCCC 1
XX
XX RESULT 10
XX AAD02796/c
XX ID AAD02796 standard; DNA; 36 BP.
XX XX
XX AC AAD02796;
XX XX
XX DT 31-MAY-2001 (first entry)
XX XX
XX DE Candida albicans ERG8 coding sequence amplifying primer #2.
XX KW Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;
XX KW infection; PCR primer; ss.
XX OS Candida albicans.
XX XX
XX PN WO200114533-A2.
XX XX
XX PD 01-MAR-2001.
XX XX
XX PF 15-AUG-2000; 2000WO-GB003100.
XX XX
XX PR 21-AUG-1999; 99GB-00019766.
XX XX
XX PA (ASTR ) ASTRAZENECA AB.
XX PA (ASTR ) ASTRAZENECA UK LTD.
XX PI Rosamond JDC, Schnell NF;
XX XX WPI; 2001-218441/22.
XX DR
XX XX
XX PT New polypeptides and polynucleotides (ERG8) from Candida albicans, useful
XX PT in assays for identifying inhibitors of phosphomevalonate kinase activity
XX PT and as reagents for diagnosing C. albicans infection.
```

XX Example 4; Page 29, 29pp; English.
PS
CC The patent discloses phosphomevalonate kinase (PMK; ERG8) protein and
CC their corresponding DNAs from *Candida albicans*. The ERG8 protein is
CC useful in assays for identifying compounds that inhibit phospho-
CC mevalonate kinase (PMK) activity. These inhibitors are useful as anti-
CC fungal agents. The ERG8 DNA and protein are also useful as reagents for
CC diagnosing *C. albicans* infection. The present DNA sequence is PCR primer
CC which is used to amplify the *Candida albicans* ERG8 coding sequence. This
CC sequence incorporates restriction enzyme sites in the ERG8 coding
CC sequence and facilitate its cloning
XX
SQ Sequence 36 BP; 9 A; 8 C; 6 G; 13 T; 0 U; 0 Other;

Query Match 1.8%; Score 24; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1276 CCAGAGACTATATAGGTTTATA 1299
Db |||||
36 CCAGAGACTATATAGGTTTATA 13

RESULT 11
ID AAS23789/c
XX AAS23789 standard; DNA; 22 BP.
AC AAS23789;
XX
XX 04-DEC-2001 (first entry)
DT
DE Primer B #42 used as probe for identifying *C. albicans* GRACE strain.
XX
XX Gene identification; essential gene; GRACE; pathogenic fungus;
KM gene replacement and conditional expression; fungal infection; probe; ss.
XX
OS *Candida albicans*.
OS Synthetic.
OS
XX WO200160975-A2.
XX
XX 23-AUG-2001.
PD
XX 20-FEB-2001; 2001WO-US005551.
PF
XX 18-FEB-2000; 2000US-0183534P.
PR
XX (ELIT-) ELITRA PHARM INC.
PA
XX Roemer T, Jiang B, Boone C, Bussey H;
PI
XX WPI; 2001-489080/53.
DR
XX
XX Identifying genes essential to fungal metabolisms and identifying
PT potential therapeutic agents that target these genes.
PT
XX
XX Claim 36; Page 320; 324pp; English.
PS
XX The present invention relates to novel methods for constructing fungal
CC strains useful for identification and validation of gene products as
CC targets for therapeutic agents, for creating a collection of identified
CC essential genes, and screening assays for the discovery of new drugs. The
CC invention provides the GRACB (gene replacement and conditional
CC expression) method for the construction of mutant organisms referred to
CC as GRACB strains of the organism. The invention can be applied to any
CC organism, particularly a pathogenic fungus e.g. *Candida albicans*,
CC *Aspergillus fumigatus* and *Cryptococcus neoformans*. The methods are useful
CC to identify agents that may be used in the treatment of fungal
CC infections. AAS23748-AAS23808 represent primers B #1-61 used as probes
CC for identifying *C. albicans* GRACB strains
XX
XX Sequence 22 BP; 8 A; 7 C; 3 G; 4 T; 0 U; 0 Other;

Query Match 1.7%; Score 22; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 GGATTAGTGTGACGTTTGCCA 475
Db |||||
22 GGATTAGTGTGACGTTTGCCA 1

RESULT 12
ID AB230804/c
XX AB230804 standard; DNA; 22 BP.
AC AB230804;
XX
XX 30-JAN-2003 (first entry)
DT
XX
DE *Candida albicans* GRACE strain PCR primer SEQ ID NO 5023.
DE
XX
XX Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KM signal transduction; DNA replication; cell division; growth;
KM proliferation; *Candida albicans*; fungicide; antifungal; PCR; primer; ss.
XX
OS *Candida albicans*.
OS
XX WO200253728-A2.
XX
XX 11-JUL-2002.
PD
XX
XX 26-DEC-2001; 2001WO-US049486.
PF
XX
XX 29-DEC-2000; 2000US-0259128P.
PR
XX 20-FEB-2001; 2001US-00792024.
PR
XX 22-AUG-2001; 2001US-0314050P.
PR
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
PI
XX WPI; 2002-566694/60.
DR
XX
XX Constructing strains for identifying gene products as effective targets
PT for therapeutic intervention, by inactivating in the strain one allele of
PT a gene and placing other allele of the gene under conditional expression.
PT
XX
XX Claim 36; SEQ ID NO 5023; 167pp + Sequence Listing; English.
PS
XX The invention relates to constructing (M1) a strain of diploid fungal
CC cells in which both alleles of a gene are modified, comprising modifying
CC one allele by insertion or replacement by a cassette having an
CC expressible selectable marker and modifying other allele by
CC recombination, of a promoter replacement fragment with a heterologous
CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival or growth of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of *C. albicans* cells and for
CC treating infection by *C. albicans*. The present sequence is that of a PCR
CC primer used in the method of the invention. Note: The sequence data for
CC this patent is not represented in the printed specification but is based
CC on sequence information supplied to Derwent by the European Patent Office
XX

Sequence 22 BP; 8 A; 7 C; 3 G; 4 T; 0 U; 0 Other;

Query Match 1.7%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 GGATTAGTGTCAAGTTGTTGCCA 475
DB 22 GGATTAGTGTCAAGTTGTTGCCA 1

RESULT 13

AAD02789
ID AAD02789 standard; DNA; 21 BP.

AC AAD02789;

DT 31-MAY-2001 (first entry)

DE Candida albicans ERG8 gene isolating PCR primer #1.

KW Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;
infection; PCR primer; ss.

OS Candida albicans.

PN WO200114533-A2.

PD 01-MAR-2001.

PF 15-AUG-2000; 2000WO-GB003100.

PR 21-AUG-1999; 99GB-00019766.

PA (ASTR) ASTRAZENECA AB.
(ASTR) ASTRAZENECA UK LTD.

PI Rosamond JDC, Schnell NF;

DR WPI; 2001-218441/22.

PT New polypeptides and polynucleotides (ERG8) from Candida albicans, useful
in assays for identifying inhibitors of phosphomevalonate kinase activity
and as reagents for diagnosing C. albicans infection.

PS Example 1; Page 24; 29pp; English.

CC The patent discloses phosphomevalonate kinase (PMK; ERG8) protein and
their corresponding DNAs from Candida albicans. The ERG8 protein is
useful in assays for identifying compounds that inhibit phospho-
mevalonate kinase (PMK) activity. These inhibitors are useful as anti-
fungal agents. The ERG8 DNA and protein are also useful as reagents for
diagnosing C. albicans infection. The present DNA sequence is a PCR
primer which is used in the isolation of a clone containing the Candida
albicans ERG8 gene

Sequence 21 BP; 6 A; 3 C; 7 G; 5 T; 0 U; 0 Other;

Query Match 1.6%; Score 21; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1129 GCTGTGTGATACGATGCAATA 1149
DB 1 GCTGTGTGATACGATGCAATA 21

RESULT 14

ACH30537
ID ACH30537 standard; cDNA; 476 BP.

AC ACH30537;

DT 13-OCT-2003 (first entry)

DE Human testis cDNA #923.

KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
genome mapping; biodiversity; genetic disorder.

OS Homo sapiens.

PN US2003073623-A1.

PD 17-APR-2003.

PF 30-JUL-2001; 2001US-00918995.

PR 30-JUL-2001; 2001US-00918995.

PA (DRMA/) DRMANAC R T.
(LABA/) LABAT I.
(STAC/) STACHE-CRAIN B.
(DICK/) DICKSON M C.
(JONE/) JONES L W.

PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

DR WPI; 2003-615964/58.

PT New polynucleotide sequences obtained from various cDNA libraries, useful
as hybridization probes, as oligomers for PCR, for chromosome and gene
mapping, in the recombinant production of protein, or in generating
antisense DNA or RNA.

PS Claim 1; SEQ ID NO 17749; 44pp; English.

CC The invention relates to an isolated polynucleotide comprising any one of
38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
determined by the technique of SBH (sequencing by hybridisation). Also
included is a purified polypeptide comprising a sequence corresponding to
a reading frame of the novel polynucleotide. The nucleic acid sequences
are useful in diagnostics as expressed sequence tags (EST) for
identifying expressed genes or for physical mapping of the human genome,
in forensics, in assessing biodiversity, or in identifying mutations
responsible for genetic disorders and other traits. The nucleotide
sequences are also useful as hybridisation probes, as oligomers for PCR,
for chromosome and gene mapping, in the recombinant production of
protein, or in generating antisense DNA or RNA. The purified polypeptide
is useful for generating antibodies specific for it. The present sequence
is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic format directly from USPTO at
segdata.uspto.gov/sequence.html?docID=20030073623

Sequence 476 BP; 120 A; 89 C; 139 G; 117 T; 0 U; 11 Other;

Query Match 1.6%; Score 21; DB 9; Length 476;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGGTGTGTGTG 1135
DB 126 GTGTGTTCCAGGTGTGTGTG 146

RESULT 15

ABK45175
ID ABK45175 standard; cDNA; 479 BP.

AC ABK45175;

DT 05-JUN-2002 (first entry)

DE cDNA encoding colon tumour protein, SEQ ID NO 726.

KM Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;
 KM gene; ss.
 XX
 OS Homo sapiens.
 PN WO200212328-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 31-JUL-2001; 2001WO-US024218.
 XX
 PR 03-AUG-2000; 2000US-0223283P.
 PR 28-MAR-2001; 2001US-0279763P.
 PR 29-JUN-2001; 2001US-0302051P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI King GE, Meagher MJ, Xu J, Secrist H;
 XX
 DR WPI; 2002-241739/29.
 XX
 PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
 PT for diagnosing, preventing, and treating colon cancer, and as markers for
 PT the progression of cancer.
 XX
 PS Claim 1; SEQ ID NO 726; 147bp; English.
 XX
 CC The invention relates to polynucleotides encoding colon tumour proteins.
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
 CC compositions, such as vaccines, for the diagnosis, prevention, and
 CC treatment of colon cancer. Polynucleotide sequences may be used as
 CC hybridisation probes or primers, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and
 CC proteins in tumour cells. The compositions are useful for stimulating an
 CC immune response against cancer, particularly for the immunotherapy of
 CC colon cancer, and as markers for the progression of cancer. ABK4450-
 CC ABK46237 represent coding sequences of human colon tumour proteins of the
 CC invention. Note: With the exception of SEQ ID NO 1 and 2, the sequence
 CC data for this patent did not form part of the printed specification but
 CC was supplied by the European Patent Office
 XX
 SQ Sequence 479 BP; 127 A; 93 C; 129 G; 130 T; 0 U; 0 Other;
 XX
 Query Match 1.6%; Score 21; DB 6; Length 479;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1115 GTGTGTTCCAGGTGCTGTG 1135
 Db 41 GTGTGTTCCAGGTGCTGTG 61
 XX
 RESULT 16
 ABV86952/c
 ID ABV86952 standard; cDNA; 518 BP.
 XX
 AC ABV86952;
 XX
 DT 13-DEC-2002 (first entry)
 XX
 DE Human colon cancer related cDNA SEQ ID NO 263.
 XX
 KM Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine; gene;
 KM ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200258534-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 16-NOV-2001; 2001WO-US043704.
 XX

PR 20-NOV-2000; 2000US-0252222P.
 PR 06-FEB-2001; 2001US-0267011P.
 PR 28-MAR-2001; 2001US-0279670P.
 PR 10-JUL-2001; 2001US-0304037P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Stolk JA, Chenault RA, Meagher MJ, Secrist H, King GE;
 XX
 DR WPI; 2002-608400/65.
 XX
 PT New isolated tumor colon polynucleotide and polypeptide, useful for the
 PT diagnosis, prevention and/or treatment of cancer, in particular colon
 PT cancer.
 XX
 PS Claim 1; SEQ ID NO 263; 266bp + Sequence Listing; English.
 XX
 CC The invention relates to a human colon tumour expressed polynucleotide
 CC (I) encoding a polypeptide (II), ABP67991-ABP67996 comprising: (i) any of
 CC 2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)
 CC complements of (i); (iii) at least 20 contiguous residues of (i); (iv)
 CC sequences that hybridize to (i), under moderately stringent conditions;
 CC (v) sequences having at least 75% or 90% identity to (i); or (vi)
 CC degenerate variants of (i). The compositions and methods of the present
 CC invention are useful for the diagnosis, prevention and/or treatment of
 CC cancer, particularly colon cancer. (I) can be used in gene therapy and
 CC (i) and (ii) are useful in pharmaceutical compositions such as vaccines.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 518 BP; 135 A; 133 C; 98 G; 145 T; 0 U; 7 Other;
 XX
 Query Match 1.6%; Score 21; DB 6; Length 518;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1115 GTGTGTTCCAGGTGCTGTG 1135
 Db 414 GTGTGTTCCAGGTGCTGTG 394
 XX
 RESULT 17
 ABK45703/c
 ID ABK45703 standard; cDNA; 530 BP.
 XX
 AC ABK45703;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE cDNA encoding colon tumour protein, SEQ ID NO 1254.
 XX
 KM Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;
 KM gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200212328-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 31-JUL-2001; 2001WO-US024218.
 XX
 PR 03-AUG-2000; 2000US-0223283P.
 PR 28-MAR-2001; 2001US-0279763P.
 PR 29-JUN-2001; 2001US-0302051P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI King GE, Meagher MJ, Xu J, Secrist H;
 XX
 DR WPI; 2002-241739/29.
 XX

PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT for diagnosing, preventing, and treating colon cancer, and as markers for
XX the progression of cancer.
PS Claim 1; SEQ ID NO 1254; 147pp; English.
XX
CC The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
CC compositions, such as vaccines, for the diagnosis, prevention, and
CC treatment of colon cancer. Polynucleotide sequences may be used as
CC hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. The compositions are useful for stimulating an
CC immune response against cancer, particularly for the immunotherapy of
CC colon cancer, and as markers for the progression of cancer. ABK4450-
CC ABK46237 represent coding sequences of human colon tumour proteins of the
CC invention. Note: With the exception of SEQ ID NO 1 and 2, the sequence
CC data for this patent did not form part of the printed specification but
XX was supplied by the European Patent Office
SQ Sequence 530 BP; 137 A; 139 C; 105 G; 149 T; 0 U; 0 Other;
Query Match 1.6%; Score 21; DB 6; Length 530;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGGGTTCCAGGCTGCTGGTG 1135
DB 439 GTGGGTTCCAGGCTGCTGGTG 419
RESULT 18
AA197097
ID AA197097 standard; cDNA; 771 BP.
XX
AC AA197097;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 3172.
XX
DE Human neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS Homo sapiens.
XX
PN WO200166719-A1.
XX
PD 13-SEP-2001.
XX
PF 02-MAR-2001; 2001WO-JP001629.
XX
PR 07-MAR-2000; 2000JP-00159195.
XX
PA (CHIB-) CHIBA PREFECTURE.
XX
PA (HISM) HISAMITSU PHARM CO LTD.
XX
XX
PI Nakagawara A;
XX
PI
XX
DR WPI, 2001-565584/63.
XX
XX Nucleic acids originating in gene expressed in human neuroblastoma,
PT useful as probe or primer in diagnosing prognosis of human neuroblastoma,
PT malignancy and susceptibility indicator or tumor marker for anti-cancer
XX agents.
PS Claim 1; Page 2308; 2979pp; Japanese.
XX
CC The invention relates to novel genes (AA19326-AA197963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes

XX
SQ Sequence 771 BP; 211 A; 143 C; 195 G; 207 T; 0 U; 15 Other;
Query Match 1.6%; Score 21; DB 4; Length 771;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGGGTTCCAGGCTGCTGGTG 1135
DB 436 GTGGGTTCCAGGCTGCTGGTG 456
RESULT 19
AB054714
ID AB054714 standard; cDNA; 1574 BP.
XX
AC AB054714;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HMSOH12 cDNA, SEQ ID NO:594.
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder; infection;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2002006771-A1.
XX
PD 03-JAN-2002.
XX
PF 07-JUN-2001; 2001WO-US018569.
XX
PR 07-JUN-2000; 2000US-0209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PA Birse CE, Rosen CA;
XX
PI WPI; 2002-147878/19.
XX
PI P-PSDB; ABP41637.
XX
DR
XX
DR WPI; 2002-147878/19.
XX
DR P-PSDB; ABP41637.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.
PS Claim 1; SEQ ID NO 594; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (AB054131-AB056105), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovarian cancer and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),

CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

SO Sequence 1574 BP; 501 A; 251 C; 359 G; 455 T; 0 U; 8 Other;

Query Match 1.6%; Score 21; DB 6; Length 1574;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGGGTTCCAGGCTGCTGGT 1135
 Db 701 GTGGGTTCCAGGCTGCTGGT 721

RESULT 20
 ID ADQ22869 standard, DNA, 1825 BP.
 AC ADQ22869;
 XX 26-AUG-2004 (first entry)
 DT 26-AUG-2004 (first entry)
 DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5689.
 XX
 KM soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
 KW da.
 OS Homo sapiens.
 XX
 PN WO2004048938-A2.
 XX
 PD 10-JUN-2004.
 XX
 PF 26-NOV-2003; 2003WO-US038193.
 XX
 PR 26-NOV-2002; 2002US-0429739P.
 XX
 PA (PROT-) PROTEIN DESIGN LABS INC.
 XX
 PI Aziz N, Ginsburg WM, Zlotnick A;
 XX
 DR WPI; 2004-441208/41.
 XX
 PT Early detection of soft tissue sarcoma comprises determining expression
 PT of a gene in a first soft tissue sample and a normal soft tissue sample
 PT and comparing the gene expression, also useful in treating soft tissue
 PT sarcoma.
 XX
 PS Example 2; SEQ ID NO 5689; 210pp; English.
 XX
 CC The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC expression of the gene in both soft tissue samples, where a higher level
 CC of protein expression in the first soft tissue sample indicates the
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cytostatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC DNA of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.
 XX

SO Sequence 1825 BP; 500 A; 379 C; 415 G; 531 T; 0 U; 0 Other;

Query Match 1.6%; Score 21; DB 12; Length 1825;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGGGTTCCAGGCTGCTGGT 1135
 Db 448 GTGGGTTCCAGGCTGCTGGT 468

RESULT 21
 ID ABV78079 standard, DNA; 2010 BP.
 XX
 AC ABV78079;
 XX
 DT 12-NOV-2002 (first entry)

DE Hypoxia-regulated protein coding sequence #99.
 XX
 XX

KM Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
 KM antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;
 KM hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
 KM ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
 KM preeclampsia; atherosclerosis; inflammatory condition; wound healing;
 KM inflammation; erythropoiesis; hair loss; human; gene; ds.

OS Homo sapiens.

PN WO200246465-A2.

PD 13-JUN-2002.

PF 10-DEC-2001; 2001WO-GB005458.

PR 08-DEC-2000; 2000GB-00030076.

PR 08-FEB-2001; 2001GB-00003156.

PR 25-OCT-2001; 2001GB-00025666.

PA (OXFO-) OXFORD BIOMEDICA UK LTD.

PI White J, Mundy CR, Ward NR, Krige D, Kingman SM, Harris RA;
 PI Rayner WM;
 XX
 DR WPI; 2002-627238/67.

PT Identifying a gene involved in disease for treating hypoxia-regulated
 PT conditions, comprises comparing the transcriptome/protome of two cell
 PT types under different conditions and identifying a differentially
 PT regulated gene;
 XX
 PS Claim 37; Page 403-404; 538pp; English.

CC The present invention relates to methods for identifying genes and
 CC proteins that are implicated in a specific disease or physiological
 CC condition. The method comprises comparing the transcriptome/protome of a
 CC specialised cell type implicated in a disease or condition with that of a
 CC second specialised cell type, under two experimental conditions, and
 CC identifying a gene that is differentially regulated in the two
 CC specialised cell types under experimental conditions. ABV7873-ABV78116
 CC and ABP65061-ABP65257 were identified using the methods of the invention.
 CC The coding sequences and proteins are useful for treating hypoxia-regulated
 CC conditions, and for manufacturing of a medicament for treating hypoxia-regulated
 CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
 CC biological response to hypoxia conditions, or hypoxic-associated
 CC pathology in a patient. The coding sequences and proteins are also useful
 CC for monitoring the therapeutic treatment of a disease or physiological
 CC condition, such as cancer, ischaemic conditions, reperfusion injury,
 CC retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory
 CC conditions, wound healing, inflammation, erythropoiesis or hair loss

XX SQ Sequence 2010 BP; 607 A; 388 C; 498 G; 517 T; 0 U; 0 Other;
 Query Match 1.6%; Score 21; DB 6; Length 2010;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1115 GTGTGTTCCAGTGCTGTG 1135
 DB 1303 GTGTGTTCCAGTGCTGTG 1323
 RESULT 22
 ADCC29939
 ID ADCC29939 standard; cDNA; 2461 BP.
 AC ADCC29939;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel cDNA sequence, SEQ ID NO:21.
 XX
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nocotropic;
 KW neuroprotective; anti-anemic; anticoagulant; thrombolytic; vulnery;
 KW anticancer; osteopathic; immunosuppressive; anti-inflammatory; cytostatic;
 KW gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029271-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002WO-US030474.
 XX
 PR 24-SEP-2001; 2001US-032463P.
 XX
 PA (HYGE-) HYSEQ INC.
 XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX
 DR WPI; 2003-371981/35.
 DR P-PSDB; ADCC30910.
 XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Claim 1; SEQ ID NO 21; 1185bp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADCC29919-
 CC ADCC30889) and the polypeptides they encode (ADCC30890-ADCC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; an antibody
 CC identifying a compound which binds to a polypeptide; and methods of
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADCC1861-ADCC32627) and the polypeptides encoded by the contigs (ADCC32628

CC -ADCC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridization probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 2461 BP; 745 A; 438 C; 604 G; 674 T; 0 U; 0 Other;
 Query Match 1.6%; Score 21; DB 10; Length 2461;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGCTGTG 1135
 DB 1168 GTGTGTTCCAGTGCTGTG 1188

RESULT 23

ADL82850
 ID ADL82850 standard; cDNA; 2562 BP.
 AC ADL82850;
 XX

DT 17-JUN-2004 (first entry)
 XX

DE Human PRO4912 cDNA, SEQ ID 52.
 XX

KW Immunosuppressive; Cytostatic; Antiarthritic; Antirheumatic; Antianemic;
 KW Antiallergic; Muscular; Neuroprotective; Nephrotropic; Antiinflammatory;
 KW Gene therapy; PRO; B cell related disorder; cancer;
 KW Immune-mediated inflammatory disease; human; gene; ss.
 XX

OS Homo sapiens.
 XX

PN WO2004024097-A2.
 XX

PD 25-MAR-2004.
 XX

PF 15-SEP-2003; 2003WO-US029097.
 XX

PR 16-SEP-2002; 2002US-0411392P.
 XX

PA (GENTH) GENENTECH INC.
 XX

PI Chiu H, Clark H, Dennis K, Fong S, Schoenfeld JR, Wood W;
 PI Wu TD;
 XX

DR WPI; 2004-329389/30.
 DR P-PSDB; ADL82851.
 XX

PT New PRO polypeptide, useful for diagnosing and treating a B cell related
 PT disorder, e.g. Burkitt's lymphoma, rheumatoid arthritis, autoimmune
 PT mediated hemolytic anemia, myasthenia gravis or ankylosing spondylitis.
 XX
 PS Claim 2; Fig 52; 695bp; English.
 XX

CC The present invention relates to PRO proteins and their coding sequences.
 CC The PRO proteins are useful for diagnosing and treating a B cell related
 CC disorder, e.g. X-linked infantile hypogammaglobulinemia, polysaccharide
 CC antigen unresponsiveness, selective Iga deficiency, selective IgM
 CC deficiency, selective deficiency of IgG subclasses, immunodeficiency with
 CC hyper IgM, transient hypogammaglobulinemia of infancy, Burkitt's

CC lymphoma, intermediate lymphoma, follicular lymphoma, type II
 CC hypergenitvity, rheumatoid arthritis, autoimmune mediated haemolytic
 CC anaemia, myasthenia gravis, hypoadrenocorticism, glomerulonephritis, or
 CC ankylosing spondylitis. The PRO proteins are also useful for preparing a
 CC medicament for treating a condition that is responsive to the PRO
 CC protein, e.g. cancer or immune-mediated inflammatory diseases. The PRO
 CC coding sequences are useful as hybridization probes in chromosome and
 CC gene mapping, in preparing PRO proteins, or in generating transgenic
 CC animals or knockout animals, which in turn are useful in the development
 CC and screening of therapeutically useful reagents.

Sequence 2562 BP; 782 A; 458 C; 623 G; 699 T; 0 U; 0 Other;

Query Match 1.6%; Score 21; DB 12; Length 2562;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGGTGCTGTG 1135
 |||||
 Db 1272 GTGTGTTCCAGGTGCTGTG 1292

RESULT 24

ID ADO19167 standard; cDNA; 2562 BP.

XX ADO19167;

DT 12-AUG-2004 (first entry)

XX Human PRO polynucleotide #51.

XX Human; PRO; gene; ss; immune related disorder;
 KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
 KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
 KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
 KW renal disease; demyelinating disease; central nervous system;
 KW peripheral nervous system; demyelinating polynuropathy;
 KW Guillain-Barre syndrome;
 KW chronic inflammatory demyelinating polynuropathy.

XX Homo sapiens.

PN WO2004043361-A2.

PD 27-MAY-2004.

PF 06-NOV-2003; 2003WO-US035268.

PR 08-NOV-2002; 2002US-0425235P.

PA (GETH) GENENTECH INC.

PI Feng S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;

PI Wood WI, Wu TD;

XX WPI; 2004-420067/39.

DR P-PSDB; ADO19168.

XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
 PT treating an immune related disorder such as systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
 PT spondyloarthritis.

XX Claim 1; SEQ ID NO 106; 1731bp; English.

XX The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune

CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the
 CC central or peripheral nervous system, demyelinating polynuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polynuropathy. This sequence represents a human PRO polynucleotide of
 CC the invention.

Sequence 2562 BP; 782 A; 458 C; 623 G; 699 T; 0 U; 0 Other;

Query Match 1.6%; Score 21; DB 12; Length 2562;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGGTGCTGTG 1135
 |||||
 Db 1272 GTGTGTTCCAGGTGCTGTG 1292

RESULT 25

ID ADO60113 standard; DNA; 2562 BP.

XX ADO60113;

DT 26-AUG-2004 (first entry)

XX Human CCR6 gene SEQ ID NO:1.

XX de; human; gene; chaperonin containing T-complex 1 subunit 6; CCR6;

KW retinoblastoma pathway; cytoskeletal.

XX Homo sapiens.

FT Key Location/Qualifiers

FT CDS 56..1651

FT /*tag= a

FT /product= "CCR6"

PN WO2004048541-A2.

PD 10-JUN-2004.

PF 24-NOV-2003; 2003WO-US037548.

PR 25-NOV-2002; 2002US-0428872P.

PA (EXEL-) EXELIXIS INC.

PI Tai AK, Song C, Ollmann MM, Gillett LA, Lickteig K;

PI WPI; 2004-441173/41.

DR P-PSDB; ADO60120.

XX Use of Chaperonin containing T-complex 1 subunit 6 (CCR6) polypeptides or

PT nucleic acids, e.g. identifying a candidate retinoblastoma (RB) pathway

PT modulating agent or modulating a RB pathway of a cell.

XX Example 1; SEQ ID NO 1; 57bp; English.

XX The invention relates to a novel use of Chaperonin containing T-complex 1
 CC subunit 6 (CCR6) polypeptides or nucleic acids for identifying a
 CC candidate retinoblastoma (RB) pathway modulating agent or for modulating
 CC a RB pathway of a cell. An RB pathway modulating agent of the invention
 CC has cytostatic activity. The CCR6 polypeptides or nucleic acids are
 CC useful for identifying a candidate RB pathway modulating agent or for
 CC modulating a RB pathway of a cell. The CCR6-modulating agents are useful
 CC as therapeutic agents for treating disorders associated with defective or
 CC impaired RB function and/or CCR6 function. CCR6 interacting proteins are
 CC also useful in detecting and providing information about the function of
 CC CCR6 proteins. The present sequence represents a CCR6 nucleic acid of the
 CC invention.

Sequence 2562 BP; 782 A; 458 C; 623 G; 699 T; 0 U; 0 Other;

Query Match 1.6%; Score 21; DB 12; Length 2562;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGTGGTTCAGGTGCTGGTG 1135
DB 1272 GTGTGGTTCAGGTGCTGGTG 1292

RESULT 26
ADO60114
ID ADO60114 standard; DNA; 2562 BP.
XX
AC ADO60114;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human CCR6 gene SEQ ID NO:2.
XX
KW de; human; gene; chaperonin containing T-complex 1 subunit 6; CCT6;
KW retinoblastoma pathway; cytosstatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 56..1651
FT /*tag= a
FT /product= "CCT6"
XX
PN WO2004048541-A2.
XX
PD 10-JUN-2004.
XX
PF 24-NOV-2003; 2003WO-US037548.
XX
PR 25-NOV-2002; 2002US-0428872P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Tai AK, Song C, Ollmann MM, Gillelt LA, Lickteig K;
XX
DR WPI: 2004-441173/41.
XX
DR P-PSDB: ADO60120.
XX
PT Use of Chaperonin containing T-complex 1 subunit 6 (CCT6) polypeptides or
PT nucleic acids, e.g. identifying a candidate retinoblastoma (RB) pathway
PT modulating agent or modulating a RB pathway of a cell.
XX
PS Disclosure; SEQ ID NO 2; 57pp; English.
XX
CC The invention relates to a novel use of Chaperonin containing T-complex 1
CC subunit 6 (CCT6) polypeptides or nucleic acids for identifying a
CC candidate retinoblastoma (RB) pathway modulating agent or for modulating
CC a RB pathway of a cell. An RB pathway modulating agent of the invention
CC has cytosstatic activity. The CCT6 polypeptides or nucleic acids are
CC useful for identifying a candidate RB pathway modulating agent or for
CC modulating a RB pathway of a cell. The CCT6-modulating agents are useful
CC as therapeutic agents for treating disorders associated with defective or
CC impaired RB function and/or CCT6 function. CCT6 interacting proteins are
CC also useful in detecting and providing information about the function of
CC CCT6 proteins. The present sequence represents a CCT6 nucleic acid of the
XX invention.
XX
SQ Sequence 2562 BP; 782 A; 458 C; 623 G; 699 T; 0 U; 0 Other;

Query Match 1.6%; Score 21; DB 12; Length 2562;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGGTTCAGGTGCTGGTG 1135
DB 1272 GTGTGGTTCAGGTGCTGGTG 1292

RESULT 27
AAFI5841
ID AAFI5841 standard; cDNA; 2594 BP.
XX
AC AAFI5841;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:276.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosstatic; cardioactive; immunomodulatory; muscular;
KW vulnery; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005988.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587513/55.
XX
DR P-PSDB: AAB56638.
XX

PT Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
PS Claim 1; Page 796; 2338pp; English.
XX

CC AAFI566 to AAFI6505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytosstatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
CC AAB57303 represent sequences used in the exemplification of the present
XX invention
XX

SQ Sequence 2594 BP; 789 A; 473 C; 628 G; 703 T; 0 U; 1 Other;
Query Match 1.6%; Score 21; DB 3; Length 2594;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGGTTCAGGTGCTGGTG 1135
DB 1298 GTGTGGTTCAGGTGCTGGTG 1318

RESULT 28
ADO60115
ID ADO60115 standard; DNA; 2647 BP.
XX

```

AC ADO60115;
XX
XX 26-AUG-2004 (first entry)
XX
XX Human CCT6 gene SEQ ID NO:3.
XX
XX de; human; gene; chaperonin containing T-complex 1 subunit 6; CCT6;
XX retinoblastoma pathway; cytosolic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 137..1732
XX /*tag= a
XX /*product= "CCT6"
XX /*transl_except= (pos:1037..1039,aa:Ser)
XX /*transl_except= (pos:1043..1045,aa:Asp)
XX
XX MO2004048541-A2.
XX
XX 10-JUN-2004.
XX
XX 24-NOV-2003; 2003MO-US037548.
XX
XX 25-NOV-2002; 2002US-0428872P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Tai AK, Song C, Ollmann MM, Gillet LA, Lickteig K;
XX WPI; 2004-441173/41.
XX P-PSDB; ADO60120.
XX
XX Use of Chaperonin containing T-complex 1 subunit 6 (CCT6) polypeptides or
XX PT nucleic acids, e.g. identifying a candidate retinoblastoma (RB) pathway
XX PT modulating agent or modulating a RB pathway of a cell.
XX
XX Disclosure; SEQ ID NO 3; 57pp; English.
XX
XX The invention relates to a novel use of Chaperonin containing T-complex 1
XX CC subunit 6 (CCT6) polypeptides or nucleic acids for identifying a
XX CC candidate retinoblastoma (RB) pathway modulating agent or for modulating
XX CC a RB pathway of a cell. An RB pathway modulating agent of the invention
XX CC has cytosolic activity. The CCT6 polypeptides or nucleic acids are
XX CC useful for identifying a candidate RB pathway modulating agent or for
XX CC modulating a RB pathway of a cell. The CCT6-modulating agents are useful
XX CC as therapeutic agents for treating disorders associated with defective or
XX CC impaired RB function and/or CCT6 function. CCT6 interacting proteins are
XX CC also useful in detecting and providing information about the function of
XX CC CCT6 proteins. The present sequence represents a CCT6 nucleic acid of the
XX CC invention.
XX
XX Sequence 2647 BP; 796 A; 487 C; 654 G; 710 T; 0 U; 0 Other;
SQ

```

```

Query Match 1.6%; Score 21; DB 12; Length 2647;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1115 GTGTGTTCCAGTGCTGTG 1135
DB 1353 GTGTGTTCCAGTGCTGTG 1373

```

```

RESULT 29
ABV24814
ID ABV24814 standard; cDNA; 2837 BP.
XX
XX ABV24814;
XX
XX 16-SBP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 24805.
XX

```

```

KW Human; prostate cancer; cytosolic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001MO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189662P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 4756; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

```

```

SQ Sequence 2837 BP; 855 A; 538 C; 680 G; 764 T; 0 U; 0 Other;

```

```

Query Match 1.6%; Score 21; DB 5; Length 2837;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1115 GTGTGTTCCAGTGCTGTG 1135
DB 1307 GTGTGTTCCAGTGCTGTG 1327

```

```

RESULT 30
AAK79798
ID AAK79798 standard; DNA; 16682 BP.
XX
XX AAK79798;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34610.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytosolic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX

```

PR 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUL-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217486P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0228343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239936P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 01-NOV-2000; 2000US-0244674P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249300P.
PR 17-NOV-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-483426/52.
XX
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,

PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 34610; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 16682 BP, 4217 A; 3754 C; 4167 G; 4538 T; 0 U; 6 Other;

Query Match 1.6%; Score 21; DB 4; Length 16682;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 CTTGCTGTGGATATTGGTT 57
DB 1920 CTTGCTGTGGATATTGGTT 1940

RESULT 31
AAK70010
ID AAK70010 standard; DNA; 16682 BP.
XX
AC AAK70010;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24822.
XX
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; de.
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-APR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235534P.
PR 27-SEP-2000; 2000US-0235636P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239357P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.

XX	Sequence	16682 BP;	4217 A;	3754 C;	4167 G;	4538 T;	0 U;	6 Other;
SO	Query Match		1.6%;	Score 21;	DB 4;	Length 16682;		
	Best Local Similarity		100.0%;	Prod. No. 12;				
	Matches	21;	Conservative	0;	Mismatches	0;	Indels	0;
QY		37	CTTGCTGGTGGATATTGGTT	57				
Db		1920	CTTGCTGGTGGATATTGGTT	1940				

ID	AAK79799	standard; DNA; 21436 BP.

DT 07-NOV-2001 (first entry)

Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34611.

AA Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 Y cytostatic; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

PN WO200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US001354

PR 31-JAN-2000; 2000US-0179065P

PR 24-FEB-2000; 2000US-0184664P

16-MAR-2000; 2000US-0189874P

PR 18-APR-2000; 2000US-0198123P

07-JUN-2000; 2000US-0209467P

30--JUN-2000: 2000US-0215135P

07-JUL-2000: 2000US-02168880P

11-JYY-2000: 2000TTS-0217496P
11-UCD-2000; 2000US-0217487P.

26-JUL-2000: 2000TIS-0320963B

28-JUL-2000; 2000US-0220964P.
14-AUG-2000: 2000US-0324519P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225214P.
14-AUG-2000; 2000US-0325255P
PB

PR
14-AUG-2000; 2000US-0225267P.
BR
14-AUG-2000; 2000US-0225268P.

14-AUG-2000; 20000US-0225270P.
14-AUG-2000; 20000US-0225270P.
14-AUG-2000; 20000US-0225270P.

14-AUG-2000; 2000US-0225757P.

14-AUG-2000; 2000US-0225759P.

22-AUG-2000; 2000US-0226681P.

22-AUG-2000; 2000US-0227182P.

30-AUG-2000; 2000US-0228924P.

01-SEP-2000; 2000US-0229343P.

01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241866P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251889P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX DR WPI; 2001-483426/52.
XX
XX PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX PS Disclosure; SEQ ID NO 34611; 3071pp + Sequence Listing; English.
XX
XX CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX CC proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX CC example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX CC that affect the activity of (I) by expressing inactive proteins or to
XX CC supplement the patients own production of (I). Additionally, (I)
XX CC polynucleotides may be used to produce the secreted (I), by inserting the
XX CC nucleic acids into a host cell and culturing the cell to express the
XX CC protein. (I) proteins and polynucleotides may be used to prevent,
XX CC diagnose and treat immune/haematopoietic-related diseases, especially
XX CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX CC to AAK87694 represent human immune/haematopoietic antigen genomic
XX CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX CC represent sequences used in the exemplification of the present invention
XX
XX
XX Sequence 21436 BP; 5267 A; 4950 C; 5418 G; 5801 T; 0 U; 0 Other;
XX
XX
XX Query Match 1.6%; Score 21; DB 4; Length 21436;
XX Best Local Similarity 100.0%; Pred No.12;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 37 CTTGCTGTGGATATTGGTT 57
XX ||||||||||||||||
XX Db 1920 CTTGCTGTGGATATTGGTT 1940
XX
XX
XX RESULT 33
XX AAK70011
XX ID AAK70011 standard; DNA; 21436 BP.
XX
XX AC AAK70011;
XX
XX XX
XX DT 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24823.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
PN WO200157182-A2.
PD 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US001354.
PF
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189887P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216888P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241825P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249292P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251866P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.

PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0255678P.
XX
XX (HUMA-) HUMMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 24823; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
XX
SQ Sequence 21436 BP; 5267 A; 4950 C; 5418 G; 5801 T; 0 U; 0 Other;

Query Match 1.6%; Score 21; DB 4; Length 21436;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 CTGCTGCTGATATTGTT 57
Db 1920 CTGCTGCTGATATTGTT 1940

RESULT 34
AAD02795
ID AAD02795 standard; DNA; 33 BP.
XX
XX AAD02795;
XX
XX 31-MAY-2001 (first entry)
XX
XX Candida albicans ERG8 coding sequence amplifying primer #1.
XX
XX Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;
KM infection; PCR primer; ss.
XX
XX Candida albicans.
OS
XX
XX WO200114533-A2.
PN
XX
XX 01-MAR-2001.
PD
XX
XX 15-AUG-2000; 2000WO-GB003100.
PF
XX
XX 21-AUG-1999; 99GB-00019766.
PR
XX
XX (ASTR) ASTRAZENECA AB.
PA (ASTR) ASTRAZENECA UK LTD.
XX
XX Rosamond JDC, Schnell NF;
PI
XX
XX WPI; 2001-218441/22.
XX
XX

PT New polypeptides and polynucleotides (ERG8) from Candida albicans, useful
PT in assays for identifying inhibitors of phosphomevalonate kinase activity
PT and as reagents for diagnosing C. albicans infection.
XX
XX
XX Example 4; Page 29; 29pp; English.
PS
XX
XX The patent discloses phosphomevalonate kinase (PMK; ERG8) protein and
CC their corresponding DNAs from Candida albicans. The ERG8 protein is
CC useful in assays for identifying compounds that inhibit phospho-
CC mevalonate kinase (PMK) activity. These inhibitors are useful as anti-
CC fungal agents. The ERG8 DNA and protein are also useful as reagents for
CC diagnosing C. albicans infection. The present DNA sequence is PCR primer
CC which is used to amplify the Candida albicans ERG8 coding sequence. This
CC sequence incorporates restriction enzyme sites in the ERG8 coding
CC sequence and facilitate its cloning
XX
XX
SQ Sequence 33 BP; 10 A; 8 C; 7 G; 8 T; 0 U; 0 Other;

Query Match 1.5%; Score 20; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTAGTGC 20
Db 14 ATGTCAAAAGCATTAGTGC 33

RESULT 35
ACH70211/c
ID ACH70211 standard; DNA; 502 BP.
XX
XX ACH70211;
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Human genome derived single exon probe #3406.
DE
XX
XX Human; probe: ss; gene expression; single exon probe; microarray;
KM alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
OS
XX
XX US2003194704-A1.
PN
XX
XX 16-OCT-2003.
PD
XX
XX 03-APR-2002; 2002US-00029386.
PF
XX
XX 03-APR-2002; 2002US-00029386.
PR
XX
XX (PENN/) PENN S G.
PA (PANK/) PANK D R.
XX (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK,
PI
XX
XX WPI; 2004-119264/12.
DR
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
PT
XX
XX Claim 15; SEQ ID NO 3406; 80pp; English.
PS
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human

CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above). The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. Note: The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the invention. Note: The sequence data for this
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704

SQ Sequence 502 BP; 122 A; 148 C; 121 G; 111 T; 0 U; 0 Other;

Query Match 1.5%; Score 20; DB 12; Length 502;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 723 CTGGGAATTCACACATGAAA 742
|||
95 CTGGGAATTCACACATGAAA 76

RESULT 36

AA181760
ID AA181760 standard; cDNA; 541 BP.

AC AA181760;

XX 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 1820.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

XX WO200164835-A2.

PN 07-SEP-2001.

PD 26-FEB-2001; 2001WO-US004927.

PR 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

XX P-PSDB; AAC01829.

PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 1820; 1399bp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 541 BP; 151 A; 63 C; 71 G; 110 T; 0 U; 146 Other;

Query Match 1.5%; Score 20; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1019 TACAAGCATTACACAAAA 1038
|||
Db 479 TACAAGCATTACACAAAA 498

RESULT 37
AB212796
ID AB212796 standard; DNA; 630 BP.

XX AB212796;

XX 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 601.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

XX WO200216655-A2.

PN 28-FEB-2002.

PD 24-AUG-2001; 2001WO-US026685.

PR 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

PA (SCRI) SCRIPES RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

DR WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 601; 577bp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array of probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that

CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 630 BP, 162 A, 136 C, 142 G, 190 T, 0 U, 0 Other;

Query Match 1.5%; Score 20; DB 6; Length 630;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 CTGCTGATATTGTTCTT 60
|||||
DB 317 CTGCTGATATTGTTCTT 336

RESULT 38

AB076676/c
ID AB076676 standard; DNA; 46852 BP.

AC AB076676;

XX 26-MAR-2003 (first entry)

XX Androgen receptor signalling pathway-associated DNA AH007803.

XX Androgen receptor; transactivation; modulator; Smad3; Smad4; Akt; TGF-B;

XX signal transduction pathway; transforming growth factor-B; phosphatase;

XX tennin; cytoskeletal; antiproliferative; cellular proliferation; cancer;

XX AH007803; ds.

XX Synthetic.

XX WO200282081-A2.

XX 17-OCT-2002.

XX 05-APR-2002; 2002WO-US011086.

XX 06-APR-2001; 2001US-0282266P.

XX 13-MAR-2002; 2002US-0365060P.

XX (UVRP) UNIV ROCHESTER.

XX Chang C;

XX WPI; 2003-046871/04.

XX Modulating androgen receptor activity, by administering a compound that

XX PT modulates receptor activity, inhibits receptor-signal transduction

XX PT pathway/receptor-coactivator interaction or changes amount or receptor.

XX Disclosure; Page 227-240; 302pp; English.

XX This invention describes a novel method for modulating androgen receptor

XX activity or androgen receptor-mediated transactivation activity in a

XX cell. The method involves administering a compound which causes

XX modulation of the androgen receptors activity and the inhibition of

XX interaction between the receptor and a protein involved in a signal

XX transduction pathway. The compound also inhibits the interaction between

XX the androgen receptor and a protein selected from Smad3, Smad4, Akt,

XX transforming growth factor (TGF)-B and phosphatase and tennin homologues

XX deleted on chromosome 10 (PTEN) or their fragments. The compounds of the

Best Local Similarity 100.0%; Pred. No. 36;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 139 AAAGATCTAGATCAAAAT 158
|||||
DB 17354 AAAGATCTAGATCAAAAT 17335

Query Match 1.5%; Score 20; DB 8; Length 218336;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 AAAGATCTAGATCAAAAT 158
|||||
DB 82361 AAAGATCTAGATCAAAAT 82342

RESULT 39

AB076678/c
ID AB076678 standard; DNA; 218336 BP.

AC AB076678;

XX 26-MAR-2003 (first entry)

XX Androgen receptor signalling pathway-associated DNA AF067844.

XX Androgen receptor; transactivation; modulator; Smad3; Smad4; Akt; TGF-B;

XX signal transduction pathway; transforming growth factor-B; phosphatase;

XX tennin; cytoskeletal; antiproliferative; cellular proliferation; cancer;

XX AF067844; ds.

XX Synthetic.

XX WO200282081-A2.

XX 17-OCT-2002.

XX 05-APR-2002; 2002WO-US011086.

XX 06-APR-2001; 2001US-0282266P.

XX 13-MAR-2002; 2002US-0365060P.

XX (UVRP) UNIV ROCHESTER.

XX Chang C;

XX WPI; 2003-046871/04.

XX Modulating androgen receptor activity, by administering a compound that

XX PT modulates receptor activity, inhibits receptor-signal transduction

XX PT pathway/receptor-coactivator interaction or changes amount or receptor.

XX Disclosure; Page 241-299; 302pp; English.

XX This invention describes a novel method for modulating androgen receptor

XX activity or androgen receptor-mediated transactivation activity in a

XX cell. The method involves administering a compound which causes

XX modulation of the androgen receptors activity and the inhibition of

XX interaction between the receptor and a protein involved in a signal

XX transduction pathway. The compound also inhibits the interaction between

XX the androgen receptor and a protein selected from Smad3, Smad4, Akt,

XX transforming growth factor (TGF)-B and phosphatase and tennin homologues

XX deleted on chromosome 10 (PTEN) or their fragments. The compounds of the

RESULT 40

ABL88534
ID ABL88534 standard; DNA; 355 BP.

XX
AC ABL88534;

XX
DT 24-MAY-2002 (first entry)

XX
DE C marmoreus mu-conopeptide Mr3-3 coding sequence.

XX
KM Mu-conopeptide; snail; venom; voltage-gated sodium channel; analgesic;

XX
KM anaesthetic; neuromuscular blocking agent; neuroprotective; pain;

XX
KM cerebroprotective; anticonvulsant; antiaging; antidiabetic;

XX
KM cardiovascular; vasotropic; cardiac; tranquilizer; antimigraine;

XX
KM neurodegenerative disease; neuromuscular disorder; gene; ds.

XX
OS Conus marmoreus.

XX
PN W0200207678-A2.

XX
PD 31-JAN-2002.

XX
PF 23-JUL-2001; 2001WO-US023125.

XX
PR 21-JUL-2000; 2000US-0219619P.

XX
PR 03-NOV-2000; 2000US-0245157P.

XX
PR 29-JAN-2001; 2001US-0264319P.

XX
PR 21-MAR-2001; 2001US-0277270P.

XX
PA (UTAH) UNIV UTAH RES FOUND.

XX
PI (COGN-) COGNETIX INC.

XX
PI Oliveira BM, McIntosh JM, Garrett JE, Watkins M, Cruz LJ, Shon K;

XX
PI Jacobsen R, Jones RM, Cartier GE, Shen GS;

XX
DR WPI; 2002-217020/27.

XX
DR P-PSDB; ABB88246.

XX
PT New mu-conopeptides useful for treating disorders associated with voltage

XX
PT agents, as local anesthetic agents, as analgesic agents and as

XX
PT neuroprotective agents.

XX
PS Claim 8; Page 48; 231pp; English.

XX
CC The present invention relates to mu-conopeptides derived from snails,

XX
CC which can be in the treatment of disorders associated with voltage-gated

XX
CC ion channels. These may include neurodegenerative disorders such as

XX
CC amyotrophic lateral sclerosis, neuromuscular disorders, post-operative or

XX
CC severe chronic pain, stroke, trigeminal neuralgia, diabetic neuropathy,

XX
CC post-herpetic neuralgia, neuronal pain and phantom limb, burn pain,

XX
CC epilepsy, for reducing neurotoxic injury associated with hypoxia, anoxia

XX
CC or ischemia which typically follows stroke, cerebrovascular accident,

XX
CC brain or spinal cord trauma, myocardial infarct, physical trauma,

XX
CC drowning, suffocation, perinatal asphyxia, or hypoglycaemic events, for

XX
CC providing muscle relaxation, treating essential blepharospasm and other

XX
CC forms of focal dystonia, and for anti-wrinkle use. The present sequence

XX
SQ is a mu-conopeptide preprotein coding sequence of the invention

XX
SQ Sequence 355 BP; 72 A; 93 C; 91 G; 99 T; 0 U; 0 Other;

XX
SQ Query Match 1.5%; Score 19; DB 6; Length 355;

XX
SQ Best Local Similarity 100.0%; Pred. No. 1.3e+02;

XX
SQ Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
QY 861 TGTGTGATGACCGCTT 879

XX
Db 262 TGTGTGATGACCGCTT 280

XX
Search completed: January 26, 2005, 08:04:55

XX
Job time : 536 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 09:09:55 ; Search time 552 Seconds
(without alignments)
13521.574 Million cell updates/sec

Title: US-10-069-062-6

Perfect score: 1299

Sequence: 1 atgcctaaagatttagtgc.....aagactatagatttataa 1299

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 4300275 seqs, 2872944193 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	891	68.6	1299	US-10-033-585-6023	Sequence 6023, Ap
2	85	5.0	65	US-10-033-585-1023	Sequence 3023, Ap
3	22	1.7	22	US-10-033-585-5023	Sequence 5023, Ap
4	22	1.7	690	US-10-027-632-150718	Sequence 150718, Ap
5	22	1.7	690	US-10-027-632-150718	Sequence 150718, Ap
6	22	1.7	711	US-10-027-632-26202	Sequence 26202, A
7	22	1.7	711	US-10-027-632-26202	Sequence 26202, A
8	22	1.7	711	US-10-027-632-26202	Sequence 26202, A
9	22	1.7	711	US-10-027-632-26202	Sequence 26202, A
10	21	1.6	313	US-10-242-535A-6452	Sequence 6452, Ap
11	21	1.6	313	US-10-085-783A-6452	Sequence 6452, Ap
12	21	1.6	400	US-10-767-795-5310	Sequence 5310, Ap

13	21	1.6	476	US-09-918-995-17749	Sequence 17749, A
14	21	1.6	479	US-09-920-300A-726	Sequence 726, Ap
15	21	1.6	479	US-10-033-528-726	Sequence 726, Ap
16	21	1.6	479	US-10-099-926-726	Sequence 726, Ap
17	21	1.6	518	US-09-998-598-263	Sequence 263, Ap
18	21	1.6	530	US-09-920-300A-1254	Sequence 1254, Ap
19	21	1.6	530	US-10-033-528-1254	Sequence 1254, Ap
20	21	1.6	530	US-10-099-926-1254	Sequence 1254, Ap
21	21	1.6	1060	US-10-198-846-10192	Sequence 10192, A
22	21	1.6	1574	US-10-264-049-594	Sequence 594, Ap
23	21	1.6	1825	US-10-722-860-5689	Sequence 5689, Ap
24	21	1.6	2010	US-10-170-385-414	Sequence 414, Ap
25	21	1.6	2010	US-10-172-118-662	Sequence 662, Ap
26	21	1.6	2010	US-10-159-553-993	Sequence 393, Ap
27	21	1.6	2010	US-10-342-887-662	Sequence 662, Ap
28	21	1.6	2594	US-09-925-300-276	Sequence 276, Ap
29	21	1.6	2837	US-10-357-930-24803	Sequence 24803, A
30	20	1.5	371	US-09-796-692-9039	Sequence 9039, Ap
31	20	1.5	371	US-10-040-862-9039	Sequence 9039, Ap
32	20	1.5	371	US-10-057-475B-9039	Sequence 9039, Ap
33	20	1.5	371	US-10-154-884B-9039	Sequence 9039, Ap
34	20	1.5	371	US-10-764-374-9039	Sequence 3406, Ap
35	20	1.5	502	US-10-029-386-3406	Sequence 3406, Ap
36	20	1.5	630	US-09-938-842A-601	Sequence 601, Ap
37	20	1.5	630	US-09-938-842A-601	Sequence 601, Ap
38	20	1.5	728	US-10-027-632-160412	Sequence 160412, A
39	20	1.5	728	US-10-027-632-160412	Sequence 160412, A
40	20	1.5	1656	US-10-369-483-43499	Sequence 43499, A
41	20	1.5	46852	US-10-473-939-8	Sequence 8, Appl
42	20	1.5	218336	US-10-473-939-10	Sequence 10, Appl
43	19	1.5	233	US-10-424-559-47771	Sequence 47771, A
44	19	1.5	287	US-10-425-115-143271	Sequence 143271, A
45	19	1.5	355	US-09-910-009A-157	Sequence 157, Ap
46	19	1.5	355	US-10-828-478-157	Sequence 157, Ap
47	19	1.5	538	US-10-027-632-91442	Sequence 91442, A
48	19	1.5	538	US-10-027-632-91443	Sequence 91443, A
49	19	1.5	538	US-10-027-632-91444	Sequence 91444, A
50	19	1.5	538	US-10-027-632-91445	Sequence 91445, A
51	19	1.5	538	US-10-027-632-91446	Sequence 91446, A
52	19	1.5	538	US-10-027-632-91447	Sequence 91447, A
53	19	1.5	538	US-10-027-632-91448	Sequence 91448, A
54	19	1.5	538	US-10-027-632-91449	Sequence 91449, A
55	19	1.5	542	US-10-027-632-291867	Sequence 291867, A
56	19	1.5	542	US-10-027-632-291868	Sequence 291868, A
57	19	1.5	542	US-10-027-632-291869	Sequence 291869, A
58	19	1.5	542	US-10-027-632-291870	Sequence 291870, A
59	19	1.5	593	US-10-313-542-59	Sequence 89, Appl
60	19	1.5	628	US-10-027-632-270097	Sequence 270097, A
61	19	1.5	639	US-10-027-632-261652	Sequence 261652, A
62	19	1.5	639	US-10-027-632-261652	Sequence 261652, A
63	19	1.5	639	US-10-027-632-261652	Sequence 261652, A
64	19	1.5	662	US-10-027-632-237370	Sequence 237370, A
65	19	1.5	662	US-10-027-632-237371	Sequence 237371, A
66	19	1.5	662	US-10-027-632-237370	Sequence 237370, A
67	19	1.5	662	US-10-027-632-237371	Sequence 237371, A
68	19	1.5	714	US-10-767-701-19874	Sequence 19874, A
69	19	1.5	842	US-10-027-632-162890	Sequence 162890, A
70	19	1.5	842	US-10-027-632-162891	Sequence 162891, A
71	19	1.5	842	US-10-027-632-162892	Sequence 162892, A
72	19	1.5	842	US-10-027-632-162893	Sequence 162893, A
73	19	1.5	842	US-10-027-632-162891	Sequence 162891, A
74	19	1.5	842	US-10-027-632-162892	Sequence 162892, A
75	19	1.5	858	US-10-027-632-10012	Sequence 10012, A
76	19	1.5	858	US-10-027-632-10012	Sequence 10012, A
77	19	1.5	870	US-10-369-493-11994	Sequence 41994, A
78	19	1.5	894	US-10-437-963-37487	Sequence 37487, A
79	19	1.5	995	US-10-363-345A-2114	Sequence 2114, Ap
80	19	1.5	995	US-10-363-345A-2114	Sequence 2114, Ap
81	19	1.5	1111	US-10-425-114-10780	Sequence 10780, A
82	19	1.5	1140	US-10-425-115-182260	Sequence 280260, A
83	19	1.5	1157	US-10-425-114-28009	Sequence 28009, A
84	19	1.5	1231	US-10-437-963-37483	Sequence 37483, A
85	19	1.5	1275	US-10-425-115-87852	Sequence 87852, A

816	16	1.2	470	16	US-10-242-535A-43372	Sequence 43372, A
817	16	1.2	470	16	US-10-085-783A-43372	Sequence 43372, A
C 818	16	1.2	471	13	US-10-027-633-75228	Sequence 75228, A
C 819	16	1.2	471	15	US-10-027-633-75228	Sequence 75228, A
C 820	16	1.2	472	9	US-09-764-869-482	Sequence 482, App
C 821	16	1.2	472	14	US-10-091-504-482	Sequence 482, App
C 822	16	1.2	472	16	US-10-227-577-77222	Sequence 77222, A
823	16	1.2	473	13	US-10-027-633-77222	Sequence 77222, A
824	16	1.2	473	13	US-10-027-633-301557	Sequence 301557, A
825	16	1.2	473	15	US-10-027-633-77222	Sequence 77222, A
826	16	1.2	473	15	US-10-027-633-301557	Sequence 301557, A
827	16	1.2	474	10	US-09-918-995-19944	Sequence 19944, A
828	16	1.2	474	16	US-10-424-599-3886	Sequence 3886, App
C 829	16	1.2	475	16	US-10-424-599-128478	Sequence 128478, A
C 830	16	1.2	476	9	US-09-964-824A-331	Sequence 331, App
831	16	1.2	476	10	US-09-918-995-14996	Sequence 14996, A
832	16	1.2	478	9	US-09-864-761-2850	Sequence 2850, App
C 833	16	1.2	479	9	US-09-864-761-1450	Sequence 1450, App
834	16	1.2	479	9	US-09-864-761-4010	Sequence 4010, App
C 835	16	1.2	479	16	US-10-424-599-140869	Sequence 140869, A
836	16	1.2	481	17	US-10-767-701-22537	Sequence 22537, A
837	16	1.2	481	18	US-10-425-115-128770	Sequence 128770, A
838	16	1.2	483	17	US-10-021-323-6667	Sequence 6667, App
C 839	16	1.2	485	13	US-10-027-633-260594	Sequence 260594, A
C 840	16	1.2	485	15	US-10-027-633-260594	Sequence 260594, A
841	16	1.2	487	10	US-09-918-995-31734	Sequence 31734, A
C 842	16	1.2	489	9	US-09-864-761-1563	Sequence 1563, App
C 843	16	1.2	492	9	US-09-864-761-6489	Sequence 6489, App
C 844	16	1.2	493	9	US-09-864-761-15697	Sequence 15697, A
C 845	16	1.2	494	13	US-10-027-633-191056	Sequence 191056, A
C 846	16	1.2	494	15	US-10-027-633-191056	Sequence 191056, A
C 847	16	1.2	496	17	US-10-767-701-22461	Sequence 22461, A
C 848	16	1.2	498	16	US-10-242-535A-16418	Sequence 16418, A
849	16	1.2	498	16	US-10-085-783A-16418	Sequence 16418, A
850	16	1.2	499	16	US-10-242-535A-56975	Sequence 56975, A
851	16	1.2	499	16	US-10-085-783A-56975	Sequence 56975, A
852	16	1.2	501	16	US-10-424-599-13619	Sequence 13619, A
853	16	1.2	501	16	US-10-424-599-71727	Sequence 71727, A
C 854	16	1.2	502	18	US-10-856-499-1316	Sequence 1316, App
855	16	1.2	508	16	US-10-424-599-31456	Sequence 31456, A
856	16	1.2	510	9	US-09-796-692-9149	Sequence 9149, App
857	16	1.2	510	16	US-10-040-862-9149	Sequence 9149, App
858	16	1.2	510	16	US-10-057-475B-9149	Sequence 9149, App
859	16	1.2	510	16	US-10-154-884B-9149	Sequence 9149, App
860	16	1.2	510	17	US-10-115-635-354	Sequence 354, App
861	16	1.2	510	17	US-10-764-324-9149	Sequence 9149, App
C 862	16	1.2	513	18	US-10-856-499-303	Sequence 303, App
C 863	16	1.2	514	18	US-10-357-930-17449	Sequence 17449, A
864	16	1.2	516	9	US-09-796-692-4518	Sequence 4518, App
865	16	1.2	516	14	US-10-040-862-4518	Sequence 4518, App
866	16	1.2	516	16	US-10-057-475B-4518	Sequence 4518, App
867	16	1.2	516	16	US-10-154-884B-4518	Sequence 4518, App
868	16	1.2	516	17	US-10-764-324-4518	Sequence 4518, App
C 869	16	1.2	518	13	US-10-027-633-77778	Sequence 77778, A
C 870	16	1.2	518	13	US-10-027-633-300597	Sequence 300597, A
871	16	1.2	518	14	US-10-247-703-38	Sequence 38, App1
872	16	1.2	518	15	US-10-029-517-101	Sequence 101, App
873	16	1.2	518	15	US-10-027-633-77778	Sequence 77778, A
C 874	16	1.2	518	15	US-10-027-633-300597	Sequence 300597, A
875	16	1.2	520	10	US-09-814-353-20689	Sequence 20689, A
876	16	1.2	520	13	US-10-027-633-205011	Sequence 205011, A
C 877	16	1.2	520	13	US-10-027-633-205012	Sequence 205012, A
C 878	16	1.2	520	15	US-10-027-633-205011	Sequence 205011, A
C 879	16	1.2	520	15	US-10-027-633-205012	Sequence 205012, A
880	16	1.2	522	11	US-09-864-408A-1565	Sequence 1565, App
881	16	1.2	523	17	US-10-021-323-12520	Sequence 12520, App
C 882	16	1.2	524	9	US-09-938-842A-4098	Sequence 4098, App
C 883	16	1.2	524	13	US-10-027-633-3073	Sequence 3073, App
C 884	16	1.2	524	13	US-10-027-633-3073	Sequence 3073, App
C 885	16	1.2	524	13	US-10-027-633-3073	Sequence 3073, App
886	16	1.2	524	15	US-10-027-633-3073	Sequence 291977, A
C 887	16	1.2	524	15	US-10-027-633-291977	Sequence 291977, A
888	16	1.2	526	9	US-09-925-299-396	Sequence 396, App

962	16	1.2	560	15	US-10-029-386-3425	Sequence 3425, Ap
963	16	1.2	560	15	US-10-027-632-93845	Sequence 93845, A
964	16	1.2	560	15	US-10-027-632-93846	Sequence 93846, A
965	16	1.2	560	15	US-10-027-632-305252	Sequence 305252, A
966	16	1.2	560	15	US-10-027-632-305253	Sequence 305253, A
967	16	1.2	561	13	US-10-027-632-290308	Sequence 290308, A
968	16	1.2	561	13	US-10-027-632-290308	Sequence 290308, A
969	16	1.2	561	18	US-10-027-632-290308	Sequence 290308, A
970	16	1.2	562	17	US-10-027-632-240337	Sequence 240337, A
971	16	1.2	564	13	US-10-027-632-240337	Sequence 240337, A
972	16	1.2	564	15	US-10-027-632-40337	Sequence 40337, A
973	16	1.2	564	15	US-10-027-632-40337	Sequence 40337, A
974	16	1.2	564	15	US-10-027-632-40337	Sequence 40337, A
975	16	1.2	565	17	US-10-027-632-4311	Sequence 4311, Ap
976	16	1.2	565	17	US-10-027-632-4311	Sequence 4311, Ap
977	16	1.2	571	9	US-09-864-761-7731	Sequence 7731, Ap
978	16	1.2	571	15	US-10-029-386-2686	Sequence 2686, Ap
979	16	1.2	572	15	US-10-029-517-18	Sequence 18, Appl
980	16	1.2	572	15	US-10-029-517-18	Sequence 5388, Ap
981	16	1.2	573	17	US-10-029-517-18	Sequence 93, Appl
982	16	1.2	573	16	US-10-424-599-66415	Sequence 66415, A
983	16	1.2	576	9	US-09-764-868-523	Sequence 523, App
984	16	1.2	576	15	US-10-029-386-4643	Sequence 4643, Ap
985	16	1.2	577	13	US-10-027-632-322970	Sequence 322970, A
986	16	1.2	577	13	US-10-027-632-322971	Sequence 322971, A
987	16	1.2	577	15	US-10-027-632-322971	Sequence 322971, A
988	16	1.2	577	15	US-10-027-632-322971	Sequence 322971, A
989	16	1.2	579	13	US-10-027-632-289790	Sequence 289790, A
990	16	1.2	580	18	US-10-027-632-289790	Sequence 289790, A
991	16	1.2	583	17	US-10-021-323-2237	Sequence 50140, A
992	16	1.2	583	17	US-10-021-323-2237	Sequence 2237, Ap
993	16	1.2	585	13	US-10-027-632-210856	Sequence 14118, A
994	16	1.2	585	13	US-10-027-632-210856	Sequence 210856, A
995	16	1.2	585	13	US-10-027-632-210856	Sequence 210856, A
996	16	1.2	585	15	US-10-027-632-210856	Sequence 210856, A
997	16	1.2	585	15	US-10-027-632-210856	Sequence 210856, A
998	16	1.2	585	15	US-10-027-632-210856	Sequence 210856, A
999	16	1.2	585	15	US-10-027-632-210856	Sequence 210856, A
1000	16	1.2	586	13	US-10-027-632-211350	Sequence 211350, A

ALIGNMENTS

RESULT 1
US-10-032-585-6023
; Sequence 6023, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jjiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patent version 3.1
; SEQ ID NO 6023
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-6023

Query Match 68.6%; Score 891; DB 15; Length 1299;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGTCAAAAGCATTTAGTGAACCTGGAAGCAATTTCTGCTGTGATATTTGCTTT 60
DB 1 ATGTCAAAAGCATTTAGTGAACCTGGAAGCAATTTCTGCTGTGATATTTGCTTT 60

QY	61	GAGCCAAATTTAGATGCTTTATGTGACAGCATTTGTCTATCAGCAATGCAATGCAATTAAACA	120
DB	61	GAGCCAAATTTAGATGCTTTATGTGACAGCATTTGTCTATCAGCAATGCAATGCAATTAAACA	120
QY	121	CCAAAAGAACCGATTTGAAAGAAATCTAGAAATCAAAATTTCTGACCCCAATTTGCAAC	180
DB	121	CCAAAAGAACCGATTTGAAAGAAATCTAGAAATCAAAATTTCTGACCCCAATTTGCAAC	180
QY	181	GGAAGATGGAAATTCACATATCATCAAAATACAGAAAGCCAGAAAGTTCACTACGC	240
DB	181	GGAAGATGGAAATTCACATATCATCAAAATACAGAAAGCCAGAAAGTTCACTACGC	240
QY	241	ATTAATTCATTTTATGAGCAACATATATATGCTTATGCTTATATCAACCGACAA	300
DB	241	ATTAATTCATTTTATGAGCAACATATATATGCTTATGCTTATATCAACCGACAA	300
QY	301	GCATTTGATCTTGAATCATATTTACTCAGACCCCTGATATCTTCAAGAAATACT	360
DB	301	GCATTTGATCTTGAATCATATTTACTCAGACCCCTGATATCTTCAAGAAATACT	360
QY	361	GAACCAAGACATCTCGAATGAGAAACAAATTTCTTACATCTGCTGCAATTAAC	420
DB	361	GAACCAAGACATCTCGAATGAGAAACAAATTTCTTACATCTGCTGCAATTAAC	420
QY	421	GAAAGTGAAGAACCGATAGTTAGTTACGCGCAGAGTTAGTGAATGTTGCCAAGT	480
DB	421	GAAAGTGAAGAACCGATAGTTAGTTACGCGCAGAGTTAGTGAATGTTGCCAAGT	480
QY	481	TTATATTCATTTTATGAGCAACATATATATGCTTATGCTTATATCAACCGAC	540
DB	481	TTATATTCATTTTATGAGCAACATATATATGCTTATGCTTATATCAACCGAC	540
QY	541	GCAAGATGACATTTGATTTATGAGCAACATATATGCTTATGCTTATATCAACCGAC	600
DB	541	GCAAGATGACATTTGATTTATGAGCAACATATATGCTTATGCTTATATCAACCGAC	600
QY	601	GCAATTTATGCTCTGATTTGATATAGAAATTTGAGCACTTTGATTAATGACGT	660
DB	601	GCAATTTATGCTCTGATTTGATATAGAAATTTGAGCACTTTGATTAATGACGT	660
QY	661	CAGTTCTGAAAGTGTCTGAGAAATTTCCCAAGATTTGAAATTTGATTAATG	720
DB	661	CAGTTCTGAAAGTGTCTGAGAAATTTCCCAAGATTTGAAATTTGATTAATG	720
QY	721	AACTGGAAATTTCAAAATGAAAGTGTATGCAATTTGATTAATGATTAATG	780
DB	721	AACTGGAAATTTCAAAATGAAAGTGTATGCAATTTGATTAATGATTAATG	780
QY	781	GACGTCAAGGTGCTGAGAAACCAACCAATTTGATTAATGATTAATGATTAATG	840
DB	781	GACGTCAAGGTGCTGAGAAACCAACCAATTTGATTAATGATTAATGATTAATG	840
QY	841	GAAGATGATCATTTCTGTGAGCTTTGACGTTGATTAATGATTAATGATTAATG	900
DB	841	GAAGATGATCATTTCTGTGAGCTTTGACGTTGATTAATGATTAATGATTAATG	900
QY	901	ATGAAGAAATTTGAGGAAATGCGTGAATAATGAGTCAAGTCAAGTCAAGTCAAGT	960
DB	901	ATGAAGAAATTTGAGGAAATGCGTGAATAATGAGTCAAGTCAAGTCAAGTCAAGT	960
QY	961	GAGTTAGATCATTTCTGTGAGCTTTGACGTTGATTAATGATTAATGATTAATG	1020
DB	961	GAGTTAGATCATTTCTGTGAGCTTTGACGTTGATTAATGATTAATGATTAATG	1020
QY	1021	CAAGCAATTTAACAATAATGAGGTTTCAATTTGAACTGATGCTCAAACTCAAGTT	1080
DB	1021	CAAGCAATTTAACAATAATGAGGTTTCAATTTGAACTGATGCTCAAACTCAAGTT	1080
QY	1081	GACCGTTGCAAGAGATTTCTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1140
DB	1081	GACCGTTGCAAGAGATTTCTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1140
QY	1141	GATGCAATGCTGTATTTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1200

Db 1141 GATGCAATAGCTGATTTAGTTGAAAAATTCAGGAAATTTTAAAGCAAACTCTT 1200
Qy 1201 GAAATCCAGATTATTTTCTAATATGTTTCTGGGTTGATTGGAGAGCAAGAGGT 1260
Db 1201 GAAATCCAGATTATTTTCTAATATGTTTCTGGGTTGATTGGAGAGCAAGAGGT 1260
Qy 1261 GTACTTGAAGAAAAACAGAGACTATATAGTTTATAA 1299
Db 1261 GTACTTGAAGAAAAACAGAGACTATATAGTTTATAA 1299
RESULT 2
US-10-032-585-3023/c
; Sequence 3023, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jjiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patent version 3.1
; SEQ ID NO 3023
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-3023

Query Match 5.0%; Score 65; DB 15; Length 65;
Best Local Similarity 100.0%; Pred. No. 2e-22;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTCAAAAGCATTTAGTGCACCTGGAAAAAGCATTTCTTGCTGATATTTGTTCTT 60
Db 65 ATGTCAAAAGCATTTAGTGCACCTGGAAAAAGCATTTCTTGCTGATATTTGTTCTT 6
Qy 61 GAGCC 65
Db 5 GAGCC 1

RESULT 3
US-10-032-585-5023/c
; Sequence 5023, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jjiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patent version 3.1
; SEQ ID NO 5023
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-5023

Query Match 1.7%; Score 22; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 454 GGATTAGTGTCAAGTTGTTGCCA 475

Db 22 GGATTAGTGTCAAGTTGTTGCCA 1
RESULT 4
US-10-027-632-150718/c
; Sequence 150718, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150718
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-150718

Query Match 1.7%; Score 22; DB 13; Length 690;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1030 ACACAAAATCAGAGGTTCCAA 1051
Db 325 ACACAAAATCAGAGGTTCCAA 304
RESULT 5
US-10-027-632-150718/c
; Sequence 150718, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150718

LENGTH: 690
TYPE: DNA
ORGANISM: Human
US-10-027-632-150718

Query Match
Best Local Similarity 1.7%; Score 22; DB 15; Length 690;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1030 ACACAAAATCAGAGTTCCAA 1051
DB 325 ACACAAAATCAGAGTTCCAA 304

RESULT 6
US-10-027-632-26202/c
Sequence 26202, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26202
LENGTH: 711
TYPE: DNA
ORGANISM: Human
US-10-027-632-26202

Query Match
Best Local Similarity 1.7%; Score 22; DB 13; Length 711;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1030 ACACAAAATCAGAGTTCCAA 1051
DB 325 ACACAAAATCAGAGTTCCAA 304

RESULT 7
US-10-027-632-26203/c
Sequence 26203, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26203
LENGTH: 711
TYPE: DNA
ORGANISM: Human
US-10-027-632-26203

Query Match
Best Local Similarity 1.7%; Score 22; DB 13; Length 711;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1030 ACACAAAATCAGAGTTCCAA 1051
DB 325 ACACAAAATCAGAGTTCCAA 304

RESULT 8
US-10-027-632-26202/c
Sequence 26202, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26202
LENGTH: 711
TYPE: DNA
ORGANISM: Human
US-10-027-632-26202

Query Match
Best Local Similarity 1.7%; Score 22; DB 15; Length 711;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1030 ACACAAAATCAGAGTTCCAA 1051
DB 325 ACACAAAATCAGAGTTCCAA 304

RESULT 9
US-10-027-632-26203/c
Sequence 26203, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.

```
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 26203
/ LENGTH: 711
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-26203
```

```
Query Match          1.6%; Score 22; DB 15; Length 711;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1030 ACACAAATCAGAGTCCAA 1051
      |||||
Db      325 ACACAAATCAGAGTCCAA 304
```

```
RESULT 10
US-10-242-535A-6452
/ Sequence 6452, Application US/10242535A
/ Publication No. US20040013663A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liev, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2005
/ CURRENT APPLICATION NUMBER: US/10/242,535A
/ PRIOR FILING DATE: 2002-09-12
/ PRIOR APPLICATION NUMBER: US 10/085,783
/ PRIOR FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 6452
/ LENGTH: 313
/ TYPE: DNA
/ ORGANISM: Human
US-10-242-535A-6452
```

```
Query Match          1.6%; Score 21; DB 16; Length 313;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1115 GTGTGTTCCAGTGTGTG 1135
      |||||
Db      36 GTGTGTTCCAGTGTGTG 56
```

RESULT 11

```
US-10-085-783A-6452
/ Sequence 6452, Application US/10085783A
/ Publication No. US20040037841A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liev, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2002
/ CURRENT APPLICATION NUMBER: US/10/085,783A
/ PRIOR FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 6452
/ LENGTH: 313
/ TYPE: DNA
/ ORGANISM: Human
US-10-085-783A-6452
```

```
Query Match          1.6%; Score 21; DB 16; Length 313;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1115 GTGTGTTCCAGTGTGTG 1135
      |||||
Db      36 GTGTGTTCCAGTGTGTG 56
```

```
RESULT 12
US-10-767-795-5310
/ Sequence 5310, Application US/10767795
/ Publication No. US20040181830A1
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Zhou, Yihua
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
/ FILE REFERENCE: 38-21(53534)B
/ CURRENT APPLICATION NUMBER: US/10/767,795
/ PRIOR FILING DATE: 2004-01-30
/ NUMBER OF SEQ ID NOS: 117596
/ SEQ ID NO 5310
/ LENGTH: 400
/ TYPE: DNA
/ ORGANISM: Gossypium hirsutum
/ FEATURE:
/ OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C61781_1
US-10-767-795-5310
```

```
Query Match          1.6%; Score 21; DB 17; Length 400;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1115 GTGTGTTCCAGTGTGTG 1135
      |||||
Db      238 GTGTGTTCCAGTGTGTG 258
```

```
RESULT 13
US-09-918-995-17749
/ Sequence 17749, Application US/09918995
/ Publication No. US20030073623A1
/ GENERAL INFORMATION:
/ APPLICANT: HySeq, Inc.
/ TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
/ TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
/ FILE REFERENCE: 20411-756
```

```

; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17749
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) ... (476)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-17749
```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 476;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1115 GTGTGTTCCAGTGTCTGTG 1135
Db 126 GTGTGTTCCAGTGTCTGTG 146
```

```

RESULT 14
US-09-920-300A-726
; Sequence 726, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 726
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-726
```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 479;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1115 GTGTGTTCCAGTGTCTGTG 1135
Db 41 GTGTGTTCCAGTGTCTGTG 61
```

```

RESULT 15
US-10-033-528-726
; Sequence 726, Application US/10033528
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 726
```

```

; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-726
```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 13; Length 479;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1115 GTGTGTTCCAGTGTCTGTG 1135
Db 41 GTGTGTTCCAGTGTCTGTG 61
```

```

RESULT 16
US-10-099-926-726
; Sequence 726, Application US/10099926
; Publication No. US2003016064A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secretist, Heather
; APPLICANT: Jians, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; CURRENT FILING DATE: 2002-03-17
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 726
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-099-926-726
```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 15; Length 479;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1115 GTGTGTTCCAGTGTCTGTG 1135
Db 41 GTGTGTTCCAGTGTCTGTG 61
```

```

RESULT 17
US-09-998-598-263/C
; Sequence 263, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jianshun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 263
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 46, 60, 64, 83, 482
; OTHER INFORMATION: n = A,T,C or G
US-09-998-598-263
```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 518;
```

Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
Db 414 GTGTGTTCCAGTGTCTGTG 394

RESULT 18
US-09-920-300A-1254/c
; Sequence 1254, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1254
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1254

Query Match 1.6%; Score 21; DB 9; Length 530;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
Db 439 GTGTGTTCCAGTGTCTGTG 419

RESULT 19
US-10-033-528-1254/c
; Sequence 1254, Application US/10033528
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1254
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-1254

Query Match 1.6%; Score 21; DB 13; Length 530;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
Db 439 GTGTGTTCCAGTGTCTGTG 419

RESULT 20
US-10-099-926-1254/c
; Sequence 1254, Application US/10099926

; Publication No. US20030166064A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; CURRENT FILING DATE: 2002-03-17
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1254
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-099-926-1254

Query Match 1.6%; Score 21; DB 15; Length 530;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
Db 439 GTGTGTTCCAGTGTCTGTG 419

RESULT 21
US-10-198-846-10192
; Sequence 10192, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10192
; LENGTH: 1060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-10192

Query Match 1.6%; Score 21; DB 14; Length 1060;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
Db 477 GTGTGTTCCAGTGTCTGTG 497

RESULT 22
US-10-264-049-594
; Sequence 594, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04

```

; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 594
; LENGTH: 1574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1467)..(1467)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1475)..(1475)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1485)..(1485)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1520)..(1520)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1550)..(1550)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1572)..(1574)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-594

Query Match
Best Local Similarity 1.6%; Score 21; DB 16; Length 1574;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTGTGTG 1135
DB 701 GTGTGTTCCAGTGTGTGTG 721

RESULT 23
US-10-723-860-5689
; Sequence 5689, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nacasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlocnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882,0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/773,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5689
; LENGTH: 1825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-5689

Query Match
Best Local Similarity 1.6%; Score 21; DB 18; Length 1825;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTGTGTG 1135
```

```

DB 448 GTGTGTTCCAGTGTGTGTG 468

RESULT 24
US-10-170-385-414
; Sequence 414, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 53682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 414
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-170-385-414

Query Match
Best Local Similarity 1.6%; Score 21; DB 15; Length 2010;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTGTGTG 1135
DB 1303 GTGTGTTCCAGTGTGTGTG 1323

RESULT 25
US-10-172-118-662
; Sequence 662, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2659
; SEQ ID NO 662
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 001762
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-662
```

Query Match 1.6%; Score 21; DB 15; Length 2010;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
DB 1303 GTGTGTTCCAGTGTCTGTG 1323

RESULT 26

US-10-159-563-393
; Sequence 393, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 393
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-393

Query Match 1.6%; Score 21; DB 16; Length 2010;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
DB 1303 GTGTGTTCCAGTGTCTGTG 1323

RESULT 27

US-10-342-887-662
; Sequence 662, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernard, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 662
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-662

Query Match 1.6%; Score 21; DB 16; Length 2010;

Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
DB 1303 GTGTGTTCCAGTGTCTGTG 1323

RESULT 28

US-09-925-300-276
; Sequence 276, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Ruben,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 276
; LENGTH: 2594
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2198)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-276

Query Match 1.6%; Score 21; DB 9; Length 2594;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
DB 1298 GTGTGTTCCAGTGTCTGTG 1318

RESULT 29

US-10-357-930-24803
; Sequence 24803, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endegge, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232

```
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 24803
/ LENGTH: 2837
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-357-930-24803

Query Match      1.6%; Score 21; DB 18; Length 2837;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1115 GTGTGTTCCAGTGTCTGTTG 1135
Db      1307 GTGTGTTCCAGTGTCTGTTG 1327

RESULT 30
US-09-796-692-9039
/ Sequence 9039, Application US/09796692
/ Publication No. US20020198362A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
/ FILE REFERENCE: 2077.001200
/ CURRENT APPLICATION NUMBER: US/09/796,692
/ PRIOR FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ NUMBER OF SEQ ID NOS: 9597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 9039
/ LENGTH: 371
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-796-692-9039

Query Match      1.5%; Score 20; DB 9; Length 371;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      723 CTGGGAATTCAACATGAAA 742
Db      139 CTGGGAATTCAACATGAAA 158

RESULT 31
US-10-040-862-9039
/ Sequence 9039, Application US/10040862
```

```
/ Publication No. US20030078396A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Ketter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: US 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR APPLICATION NUMBER: US 09/796,692
/ PRIOR FILING DATE: 2001-03-01
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 9039
/ LENGTH: 371
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-040-862-9039

Query Match      1.5%; Score 20; DB 14; Length 371;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      723 CTGGGAATTCAACATGAAA 742
Db      139 CTGGGAATTCAACATGAAA 158

RESULT 32
US-10-057-475B-9039
/ Sequence 9039, Application US/10057475B
/ Publication No. US20040002068A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Clapper, Jonathan David
/ APPLICANT: Wang, Aijun
/ APPLICANT: Ordenez, Nadia
/ APPLICANT: Carter, Lauren
/ APPLICANT: McNeill, Patricia Dianne
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-014402US
/ CURRENT APPLICATION NUMBER: US/10/057,475B
```



```
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See file Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10979
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 9039
/ LENGTH: 371
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-475B-9039
```

```
Query Match          1.5%; Score 20; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 723 CTGGGAATTCAAACATGAAA 742
|||||
Db 139 CTGGGAATTCAAACATGAAA 158
```

```
RESULT 33
US-10-154-884B-9039
/ Sequence 9039, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
/ APPLICANT: Galger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
/ PRIOR FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
```

```
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See file Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 9039
/ LENGTH: 371
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-154-884B-9039
```

```
Query Match          1.5%; Score 20; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 723 CTGGGAATTCAAACATGAAA 742
|||||
Db 139 CTGGGAATTCAAACATGAAA 158
```

```
RESULT 34
US-10-764-324-9039
/ Sequence 9039, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Galger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ PRIOR FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ Remaining Prior Application data removed - See file Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 9039
/ LENGTH: 371
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-764-324-9039
```

```
Query Match          1.5%; Score 20; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 723 CTGGGAATTCAAACATGAAA 742
|||||
Db 139 CTGGGAATTCAAACATGAAA 158
```

RESULT 35

```
US-10-029-386-3406/c
; Sequence 3406, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34286
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3406
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC015845.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
; OTHER INFORMATION: SWISSPROT HIT: Q9Y3D9, EVALUE 2.00e-29
; OTHER INFORMATION: NT HIT: AF161447.1, EVALUE 2.00e-92
; US-10-029-386-3406
; EST_HUMAN HIT: BE388985.1, EVALUE 3.00e-92
```

```
Query Match 1.5%; Score 20; DB 15; Length 502;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 723 CTGGGATTCACACATGAAA 742
Db 95 CTGGGATTCACACATGAAA 76
```

```
RESULT 36
US-09-938-842A-601
; Sequence 601, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 601
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-601
```

```
Query Match 1.5%; Score 20; DB 9; Length 630;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 41 CTGGTGATATTTGGTCTT 60
|||||
```

```
Db 317 CTGGTGATATTTGGTCTT 336
```

```
RESULT 37
US-09-938-842A-601
; Sequence 601, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 601
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-601
```

```
Query Match 1.5%; Score 20; DB 11; Length 630;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 41 CTGGTGATATTTGGTCTT 60
Db 317 CTGGTGATATTTGGTCTT 336
```

```
RESULT 38
US-10-027-632-160412/c
; Sequence 160412, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160412
; LENGTH: 728
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-160412
```

```
Query Match 1.5%; Score 20; DB 13; Length 728;
```

Best Local Similarity 100.0%; Pred. No. 33;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1117 GTGTTCCAGGTGCTGTGG 1136
Db 466 GTGTTCCAGGTGCTGTGG 447

RESULT 39
US-10-027-632-160412/c
; Sequence 160412, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 160412
; LENGTH: 728
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-160412

Query Match 1.5%; Score 20; DB 15; Length 728;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1117 GTGTTCCAGGTGCTGTGG 1136
Db 466 GTGTTCCAGGTGCTGTGG 447

RESULT 40
US-10-369-493-43499
; Sequence 43499, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianteng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 43499
; LENGTH: 1656
; TYPE: DNA
; ORGANISM: No. US2003023675A1loc punctiforme
US-10-369-493-43499

Query Match 1.5%; Score 20; DB 15; Length 1656;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 697 GAGTTGAAAAAATTGATTGA 716
Db 571 GAGTTGAAAAAATTGATTGA 590

Search completed: January 26, 2005, 11:17:55
Job time : 602 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 07:43:20 ; Search time 3060 Seconds
(without alignments)
15469.023 Million cell updates/sec

Title: US-10-069-062-6

Perfect score: 1299
Sequence: 1 atgcacaaagcatctagtc.....aagactataggtttacaa 1299

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size: 0

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

EST:
1: gb_esc1:
2: gb_esc2:
3: gb_esc3:
4: gb_esc4:
5: gb_esc5:
6: gb_esc6:
7: gb_esc7:
8: gb_esc8:
9: gb_esc9:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	2.0	730	4	BG470614 602511536
2	23	1.8	709	2	AM193927 xml19h03.x
3	23	1.8	1454	9	CL641530 CH213-13H
4	22	1.7	542	8	AQ663605 HS-3476-A
5	21	1.6	331	4	BM705592 UI-E-DWO-
6	21	1.6	349	1	AV684301 AV684301
7	21	1.6	349	1	AV696353 AV696353
8	21	1.6	349	1	AV697455 AV697455
9	21	1.6	360	1	AV695717 AV695717
10	21	1.6	368	4	BI036922 MR4-NT014
11	21	1.6	368	4	BM762828 K-EST0044
12	21	1.6	369	7	FI2095 HSC35F071.n
13	21	1.6	371	5	BP623179 BP623179
14	21	1.6	372	2	AW361182 RCL-CT025
15	21	1.6	386	6	CF141588 UI-HF-CR0
16	21	1.6	414	7	CN345987 170006000
17	21	1.6	424	7	AA150852 z144b05.r
18	21	1.6	425	7	W69312 z446g01.r1
19	21	1.6	426	5	BO347968 CM3-HT019
20	21	1.6	426	7	T66047 YC79e07.r1
21	21	1.6	434	4	BM653177 K-EST0134
22	21	1.6	434	6	CB117296 K-EST0162
23	21	1.6	441	1	AI910561 w186h07.x
24	21	1.6	443	1	AA127935 z114E07.r

25	21	1.6	459	4	BG761673
26	21	1.6	493	2	BE392663
27	21	1.6	499	2	BM641588
28	21	1.6	504	2	BE265418
29	21	1.6	506	7	CN345984
30	21	1.6	511	6	CB159634
31	21	1.6	516	4	BM796788
32	21	1.6	522	4	BG149625
33	21	1.6	528	6	CB124236
34	21	1.6	536	7	W38880
35	21	1.6	542	4	BM640031
36	21	1.6	543	6	CB119331
37	21	1.6	547	6	CB147674
38	21	1.6	548	1	AA186456
39	21	1.6	553	6	CB160975
40	21	1.6	555	8	AO837961
41	21	1.6	556	4	BM279286
42	21	1.6	561	2	BE729211
43	21	1.6	566	6	CB113018
44	21	1.6	568	2	BE543072
45	21	1.6	570	6	CB160702
46	21	1.6	576	1	AV693973
47	21	1.6	576	9	CE64965
48	21	1.6	579	4	BM640345
49	21	1.6	579	5	BQ348328
50	21	1.6	585	2	BE76343
51	21	1.6	589	4	BM641432
52	21	1.6	596	9	AG226531
53	21	1.6	597	6	CB270026
54	21	1.6	603	7	W68256
55	21	1.6	606	5	BQ348329
56	21	1.6	607	6	CA341142
57	21	1.6	609	2	AM369207
58	21	1.6	609	5	BY460801
59	21	1.6	611	6	CB124509
60	21	1.6	617	4	BM641617
61	21	1.6	623	6	CD678902
62	21	1.6	625	2	BF237622
63	21	1.6	629	7	CN345974
64	21	1.6	630	6	CD676885
65	21	1.6	632	4	BM794782
66	21	1.6	632	6	CB110558
67	21	1.6	632	7	CN345973
68	21	1.6	632	7	CN345975
69	21	1.6	635	2	BG762324
70	21	1.6	638	5	BQ347584
71	21	1.6	644	2	AA411390
72	21	1.6	648	6	CA418220
73	21	1.6	650	2	BE615514
74	21	1.6	651	2	BE535183
75	21	1.6	655	7	CN478815
76	21	1.6	658	4	BM781898
77	21	1.6	663	4	CA418049
78	21	1.6	668	6	CA419487
79	21	1.6	680	4	BI850890
80	21	1.6	680	4	BI850890
81	21	1.6	691	1	AU116915
82	21	1.6	691	2	BE390849
83	21	1.6	692	5	BQ009693
84	21	1.6	693	4	BI223761
85	21	1.6	705	2	BE251866
86	21	1.6	708	6	AU119388
87	21	1.6	708	6	CF121286
88	21	1.6	714	7	CN345956
89	21	1.6	719	6	CA439945
90	21	1.6	722	9	AG357605
91	21	1.6	722	9	AU137581
92	21	1.6	727	1	AG357605
93	21	1.6	727	4	BE383029
94	21	1.6	731	5	BO774219
95	21	1.6	731	2	BE563941
96	21	1.6	735	4	BG475173
97	21	1.6	742	1	AU126169

BG761673	602718849
BE392663	601307517
BM641588	K-EST0118
BE265418	601194958
CN345984	170006001
CB159634	K-EST0219
BM796788	K-EST0079
BG149625	nad31d12.
CB124236	K-EST0172
W38880	zb22a05.r1
BM640031	K-EST0117
CB119331	K-EST0166
CB147674	K-EST0203
AA186456	zp70h10.r
CB160975	K-EST0220
AO837961	HS_4647.05
BM279286	ad1_05
BE729211	601560832
CB113018	K-EST0154
BE543072	601069008
CB160702	K-EST0220
AV693973	AV693973
CE64965	ctgr-g88-
BM640345	K-EST0117
BQ348328	IL0-HT020
BE76343	602134224
BM641432	K-EST0118
AG226531	Lotus cor
CB270026	1008933.H
W68256	z235d05.r1
BQ348329	IL0-HT020
CA341142	haa79b03.
AM369207	IL0-HT020
BY460801	BY460801
CB124509	K-EST0173
BM641617	K-EST0118
CD678902	hg01f12.Y
BF237622	601841980
CN345974	170006001
CD676885	hoo7h03.Y
BM794782	K-EST0076
CB110558	K-EST0152
CN345973	170005326
CN345975	170004554
BG762324	602733726
BQ347584	CM4-HT019
AA411390	fh12a10.X
CA418220	UI-H-FHO-
BE615514	601278835
BE535183	601058709
CA418049	UI-CF-FNO
CA419487	UI-H-FHO-
BI850890	imgagc.1
AU116915	AU116915
BE390849	601286566
BQ009693	UI-H-ED1-
BI223761	602943145
BE251866	60115895
AU119388	AU119388
CF121286	UI-HF-CR0
CN345956	170004708
CA439945	UI-H-D10-
CA418039	170004245
AG357605	Mus muscu
AU137581	AU137581
BG383029	602764692
BO774219	UI-H-FHO-
BE563941	601348178
BG475173	602491736
AU126169	AU126169

98	21	1.6	743	4	BG285234	602409556	171	21	1.6	962	5	BQ049521	AGENCYCOURT
99	21	1.6	744	1	AL079359	DKFZP564D	172	21	1.6	969	2	BE788717	BE788717 601475806
100	21	1.6	756	4	BG421751	602450967	173	21	1.6	971	4	BG259964	BG259964 602371735
101	21	1.6	757	6	CF127733	UI-HF-ET0	174	21	1.6	979	1	AL580344	AL580344
102	21	1.6	758	4	BI087858	602852651	175	21	1.6	985	4	BM006077	BM006077 603613850
103	21	1.6	762	4	BI870861	603394694	176	21	1.6	987	4	BM465748	BM465748 AGENCYCOURT
104	21	1.6	766	2	BF685254	602415194	177	21	1.6	993	4	AL575069	AL575069
105	21	1.6	766	4	BG283849	602407577	178	21	1.6	995	9	AG129712	AG129712 Pan troy1
106	21	1.6	766	4	BG283849	602407577	179	21	1.6	1004	4	BG282819	BG282819 602405918
107	21	1.6	770	4	BG766826	6027340944	180	21	1.6	1006	1	AL579149	AL579149
108	21	1.6	774	4	BG752746	6027322772	181	21	1.6	1010	4	BM463709	BM463709
109	21	1.6	782	4	BG430826	602617767	182	21	1.6	1017	1	AL579860	AL579860
110	21	1.6	783	7	CN435921	602500321	183	21	1.6	1026	4	BG254463	BG254463 602369073
111	21	1.6	786	7	CN435921	170005321	184	21	1.6	1057	2	BF684418	BF684418 602141894
112	21	1.6	787	6	CR243315	UI-CF-FU0	185	21	1.6	1066	1	AL578909	AL578909
113	21	1.6	790	4	BG682056	602630027	186	21	1.6	1067	1	BM468576	BM468576 AGENCYCOURT
114	21	1.6	790	4	BG682056	602630027	187	21	1.6	1074	1	AL546486	AL546486
115	21	1.6	790	4	BG682056	602630027	188	21	1.6	1077	5	BU179715	BU179715 AGENCYCOURT
116	21	1.6	796	6	CA445298	UI-H-ED0-	189	21	1.6	1078	4	BM551715	BM551715 AGENCYCOURT
117	21	1.6	797	7	CF596325	AGENCYCOURT	190	21	1.6	1080	1	AL551335	AL551335
118	21	1.6	799	5	BU933792	AGENCYCOURT	191	21	1.6	1084	1	AL569497	AL569497
119	21	1.6	803	5	BM041550	603614252	192	21	1.6	1085	1	AL572419	AL572419
120	21	1.6	803	5	BU933235	AGENCYCOURT	193	21	1.6	1089	4	BM473763	BM473763
121	21	1.6	807	2	BF969343	602271604	194	21	1.6	1090	4	BM466186	BM466186
122	21	1.6	807	2	CF785928	AGENCYCOURT	195	21	1.6	1115	1	AL574833	AL574833
123	21	1.6	809	1	AU123953	AGENCYCOURT	196	21	1.6	1133	1	AL574833	AL574833
124	21	1.6	811	1	AV758159	AV758159	197	21	1.6	1147	5	BX456462	BX456462
125	21	1.6	811	4	BG611060	602612004	198	21	1.6	1153	5	BX402802	BX402802
126	21	1.6	816	4	BG339882	602438001	199	21	1.6	1154	4	BM916339	BM916339
127	21	1.6	816	6	CE601368	AGENCYCOURT	200	21	1.6	1194	5	BM902100	BM902100
128	21	1.6	823	4	BM019640	603647908	201	21	1.6	1272	3	CR594866	CR594866
129	21	1.6	833	4	BG251685	602363581	202	21	1.6	1898	3	CR613031	CR613031
130	21	1.6	833	5	BM0433203	AGENCYCOURT	203	21	1.6	1906	3	CR6009340	CR6009340
131	21	1.6	839	6	CD108078	AGENCYCOURT	204	21	1.6	1906	3	CR600067	CR600067
132	21	1.6	847	5	BX433904	EX433904	205	21	1.6	1906	3	CR612852	CR612852
133	21	1.6	854	2	BE274430	601120467	206	21	1.6	1912	3	CR626775	CR626775
134	21	1.6	855	5	BU195835	AGENCYCOURT	207	21	1.6	1917	3	CR626783	CR626783
135	21	1.6	863	5	BG689115	AGENCYCOURT	208	21	1.6	1919	3	CR605318	CR605318
136	21	1.6	865	5	BU194306	AGENCYCOURT	209	21	1.6	1948	3	CR590985	CR590985
137	21	1.6	875	4	BM040788	603614252	210	21	1.6	243	1	AU038871	AU038871
138	21	1.6	876	1	AL566339	AL566339	211	21	1.5	325	8	AA365369	AA365369
139	21	1.6	877	5	BQ650145	AGENCYCOURT	212	21	1.5	329	7	AQ097430	AQ097430
140	21	1.6	881	5	BU959564	AGENCYCOURT	213	21	1.5	366	7	BU374678	BU374678
141	21	1.6	882	5	BU959564	AGENCYCOURT	214	21	1.5	367	7	BF695474	BF695474
142	21	1.6	884	5	BO941218	AGENCYCOURT	215	21	1.5	376	2	BF694745	BF694745
143	21	1.6	891	1	AL533942	AL533942	216	21	1.5	384	2	BM47445	BM47445
144	21	1.6	892	5	BX351495	BX351495	217	21	1.5	394	1	AJ663790	AJ663790
145	21	1.6	893	5	BU172612	AGENCYCOURT	218	21	1.5	403	2	BM802365	BM802365
146	21	1.6	894	5	BO937196	AGENCYCOURT	219	21	1.5	439	8	AE019384	AE019384
147	21	1.6	898	1	AL517106	AL517106	220	21	1.5	449	5	BX568243	BX568243
148	21	1.6	900	4	BI854793	603381753	221	21	1.5	453	2	AM463092	AM463092
149	21	1.6	905	5	BX412866	BX412866	222	21	1.5	457	4	AU229270	AU229270
150	21	1.6	908	5	BO897836	AGENCYCOURT	223	21	1.5	457	4	BU398721	BU398721
151	21	1.6	916	5	BO644679	AGENCYCOURT	224	21	1.5	477	8	AQ553054	AQ553054
152	21	1.6	916	9	CNS02HR7	601445681	225	21	1.5	510	1	AV606461	AV606461
153	21	1.6	917	2	BE869765	601445681	226	21	1.5	512	9	CE001395	CE001395
154	21	1.6	927	5	BK405177	BK405177	227	21	1.5	519	2	BE808076	BE808076
155	21	1.6	928	5	BG109089	602281470	228	21	1.5	531	2	BE014089	BE014089
156	21	1.6	928	5	BU846149	AGENCYCOURT	229	21	1.5	532	2	BM480465	BM480465
157	21	1.6	931	5	BQ645262	AGENCYCOURT	230	21	1.5	543	6	CR330515	CR330515
158	21	1.6	932	2	BF686761	602140711	231	21	1.5	550	4	BG524504	BG524504
159	21	1.6	933	2	BF686761	602140711	232	21	1.5	554	4	AG129712	AG129712
160	21	1.6	935	2	BF204435	601866532	233	21	1.5	555	8	AZ429631	AZ429631
161	21	1.6	936	4	BG253821	602366513	234	21	1.5	555	8	CR315061	CR315061
162	21	1.6	939	4	BF025714	601670583	235	21	1.5	559	2	BE808336	BE808336
163	21	1.6	943	5	BO939716	AGENCYCOURT	236	21	1.5	579	2	AM257304	AM257304
164	21	1.6	947	5	BO689224	AGENCYCOURT	237	21	1.5	592	6	CA922504	CA922504
165	21	1.6	948	5	BG260794	602372379	238	21	1.5	610	2	BF639140	BF639140
166	21	1.6	950	4	BO900624	AGENCYCOURT	239	21	1.5	611	5	BU348320	BU348320
167	21	1.6	951	2	BE729363	601561563	240	21	1.5	631	5	BU957724	BU957724
168	21	1.6	951	2	BU957724	AGENCYCOURT	241	21	1.5	631	5	BU957724	BU957724
169	21	1.6	962	2	BF797060	602257908	242	21	1.5	631	5	BU957724	BU957724
170	21	1.6	962	2	BF797060	602257908	243	21	1.5	631	5	BU957724	BU957724

C 390	19	1.5	763	9	AG40581	C 463	18	1.4	183	7	CR284785
C 391	19	1.5	764	5	BU333121	464	18	1.4	183	7	CR567662
C 392	19	1.5	765	7	CO061585	465	18	1.4	194	2	BF475477
C 393	19	1.5	774	4	BI116688	466	18	1.4	198	4	BU630534
C 394	19	1.5	776	5	BO295188	467	18	1.4	202	1	AV005929
C 395	19	1.5	788	6	CA346021	468	18	1.4	206	1	AI382688
C 396	19	1.5	789	5	CC697840	469	18	1.4	209	9	CR478255
C 397	19	1.5	798	5	BM859214	470	18	1.4	213	5	BM327484
C 398	19	1.5	809	6	CF148664	471	18	1.4	216	6	CD195070
C 399	19	1.5	811	6	CA368558	472	18	1.4	225	1	AV346678
C 400	19	1.5	821	8	BZ201501	473	18	1.4	225	1	AV346678
C 401	19	1.5	825	8	BZ297872	474	18	1.4	230	8	AE742304
C 402	19	1.5	826	5	BM315978	475	18	1.4	241	7	DE0887
C 403	19	1.5	827	8	AO745023	476	18	1.4	253	8	AE058004
C 404	19	1.5	828	4	BI103926	477	18	1.4	254	1	AV083240
C 405	19	1.5	834	5	BU26259	478	18	1.4	258	4	BO843353
C 406	19	1.5	835	7	CK196698	479	18	1.4	260	4	BO3037951
C 407	19	1.5	835	7	CK196698	480	18	1.4	266	7	BB496840
C 408	19	1.5	835	7	CK196698	481	18	1.4	266	7	CO539318
C 409	19	1.5	840	9	CC720356	482	18	1.4	267	2	BB608655
C 410	19	1.5	842	6	CA119655	483	18	1.4	269	2	BB343492
C 411	19	1.5	843	8	CG949573	484	18	1.4	271	5	BM117941
C 412	19	1.5	843	8	BM592722	485	18	1.4	274	7	CR683578
C 413	19	1.5	849	5	BU291337	486	18	1.4	275	7	CR683578
C 414	19	1.5	858	7	CR291904	487	18	1.4	279	1	AV139640
C 415	19	1.5	860	9	CG969483	488	18	1.4	280	5	BM248630
C 416	19	1.5	882	2	BF262461	489	18	1.4	281	2	BF192670
C 417	19	1.5	883	9	CG448087	490	18	1.4	283	4	BG017975
C 418	19	1.5	893	9	CC702036	491	18	1.4	296	4	BG060275
C 419	19	1.5	895	9	AG141027	492	18	1.4	301	9	CL653275
C 420	19	1.5	896	5	AG141027	493	18	1.4	308	6	CD1711432
C 421	19	1.5	904	8	BM780557	494	18	1.4	309	6	CD189171
C 422	19	1.5	904	8	BM780557	495	18	1.4	312	7	F37104
C 423	19	1.5	914	5	EX433074	496	18	1.4	313	8	BZ215895
C 424	19	1.5	916	9	CG092484	497	18	1.4	316	1	AA655406
C 425	19	1.5	918	9	CG983857	498	18	1.4	317	2	BG426418
C 426	19	1.5	920	9	CNS01PMA	499	18	1.4	317	2	BM2433905
C 427	19	1.5	926	5	BM155622	500	18	1.4	319	7	TA7301
C 428	19	1.5	926	5	BM155622	501	18	1.4	322	9	CE8546705
C 429	19	1.5	929	5	BQ650256	502	18	1.4	324	9	CL579708
C 430	19	1.5	929	5	CG988749	503	18	1.4	324	9	BE242221
C 431	19	1.5	937	9	CG996006	504	18	1.4	325	2	F27990
C 432	19	1.5	940	8	AZ674704	505	18	1.4	330	8	BZ145302
C 433	19	1.5	942	8	CF412980	506	18	1.4	336	9	CR146580
C 434	19	1.5	950	7	CF412980	507	18	1.4	344	7	WM4413
C 435	19	1.5	957	9	CG367014	508	18	1.4	344	7	WM4413
C 436	19	1.5	967	9	CG382383	509	18	1.4	345	1	AA902540
C 437	19	1.5	969	5	BU37322	510	18	1.4	348	7	H80907
C 438	19	1.5	980	5	CG184068	511	18	1.4	348	7	H80907
C 439	19	1.5	987	8	CC012894	512	18	1.4	349	4	BU132623
C 440	19	1.5	1079	8	CC012894	513	18	1.4	354	1	AU144185
C 441	19	1.5	1096	9	CNS056CT	514	18	1.4	354	1	AU144185
C 442	19	1.5	1101	9	CNS056CT	515	18	1.4	354	1	AU144185
C 443	19	1.5	1105	8	CC217038	516	18	1.4	354	1	AU144185
C 444	19	1.5	1111	9	CL043555	517	18	1.4	354	1	AU144185
C 445	19	1.5	1188	9	CL118724	518	18	1.4	354	1	AU144185
C 446	19	1.5	1189	1	AL518120	519	18	1.4	354	1	AU144185
C 447	19	1.5	1206	8	CC266664	520	18	1.4	360	6	CG8114
C 448	19	1.5	1229	9	CL051231	521	18	1.4	362	1	AI269131
C 449	19	1.5	1260	9	AG127510	522	18	1.4	367	4	BG213824
C 450	19	1.5	1315	9	AG127510	523	18	1.4	367	4	BG213824
C 451	19	1.5	1459	9	CG754198	524	18	1.4	368	1	AA282120
C 452	19	1.5	1657	2	BF128477	525	18	1.4	370	6	CD171477
C 453	19	1.5	1857	9	CG756776	526	18	1.4	370	6	CD171477
C 454	19	1.5	1872	9	CG756776	527	18	1.4	374	6	CA653223
C 455	19	1.5	1872	9	CG756776	528	18	1.4	374	6	CA653223
C 456	19	1.5	1872	9	CG756776	529	18	1.4	378	1	AI952393
C 457	19	1.5	1872	9	CG756776	530	18	1.4	378	1	AI952393
C 458	19	1.5	1872	9	CG756776	531	18	1.4	378	1	AI952393
C 459	19	1.5	1872	9	CG756776	532	18	1.4	378	1	AI952393
C 460	19	1.5	1872	9	CG756776	533	18	1.4	378	1	AI952393
C 461	19	1.5	1872	9	CG756776	534	18	1.4	378	1	AI952393
C 462	19	1.5	1872	9	CG756776	535	18	1.4	378	1	AI952393

C 974	18	1.4	660	6	CA281099	SCRFSD102
C 975	18	1.4	660	8	A2974822	2M0249007
C 976	18	1.4	661	7	CN006615	CSBCS136D
C 977	18	1.4	662	5	BM114183	BM114183
C 978	18	1.4	663	2	BB640415	BB640415
C 979	18	1.4	663	6	CA299668	SCCFC16C0
C 980	18	1.4	663	6	CB053220	NISC_g113
C 981	18	1.4	663	7	CK003436	AGENCOURT
C 982	18	1.4	663	9	CC895056	ZMMBB021
C 983	18	1.4	663	9	AC150068	Pan trogl
C 984	18	1.4	664	5	BM245160	BM245160
C 985	18	1.4	664	7	CK708311	ZF201-P0
C 986	18	1.4	664	8	AZ304487	1M0004003
C 987	18	1.4	664	8	BH871928	h37f10.b
C 988	18	1.4	665	6	CA221326	SCSGFL403
C 989	18	1.4	665	6	CF201908	BR890915N
C 990	18	1.4	665	6	BM282854	BM282854
C 991	18	1.4	666	8	AO925780	RPCT-23-2
C 992	18	1.4	666	9	CC534921	CH240_413
C 993	18	1.4	667	5	BM274464	BM274464
C 994	18	1.4	667	8	BH001827	EMBA001K1
C 995	18	1.4	668	4	BM157331	f46h08.y
C 996	18	1.4	668	7	BM184567	f46g907.y
C 997	18	1.4	668	7	CN632467	taf02a04.
C 998	18	1.4	668	9	CL811621	OR_CBA002
C 999	18	1.4	669	6	CA414369	UI-H-E20-
C 1000	18	1.4	670	4	BJ803271	BJ803271

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
602511536f1 NIH_MGC_16 Homo sapiens CDNA clone IMAGE:4634169 5',
mRNA sequence.

ACCESSION
BG470614
KEYWORDS
EST.

ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: L10M1389 row: h column: 10
High quality sequence stop: 667.

FEATURES

source

1..730
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4634169"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_16"
/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAGC(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 2.0%; Score 26; DB 4; Length 730;
Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY
1110 TGGTGTGTTCCAGTGTCTGTTG 1135
|||||
DB 665 TGGTGTGTTCCAGTGTCTGTTG 690

RESULT 2
LOCUS
DEFINITION
AM193927/c 709 bp mRNA linear EST 29-NOV-1999
xm19h03.x1 NCI CGAP U-4 Homo sapiens CDNA clone IMAGE:2684693 3',
similar to SW:TCPE2_HUMAN P40227 T-COMPLEX PROTEIN 1, ZENR SUBUNIT
1, mRNA sequence.

ACCESSION
AM193927
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Seq primer: -40up from G1bco
High quality sequence stop: 405.

FEATURES

source

1..709
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2684693"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP U-4"
/note="Organ: uterus; Vector: PCWV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.48 Kb. Life Technologies catalog #:
11542-016"

ORIGIN

Query Match 1.8%; Score 23; DB 2; Length 709;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY
1113 TGGTGTGTTCCAGTGTCTGTTG 1135
|||||
DB 698 TGGTGTGTTCCAGTGTCTGTTG 676

RESULT 3
LOCUS
DEFINITION
CH213-13H21.SP6 CH213 Gastrosteus aculeatus genomic clone
CH213-13H21 3', genomic survey sequence.

ACCESSION CL641530
 VERSION CL641530.1 GI:49660954
 KEYWORDS GSS.
 SOURCE Gasterosteus aculeatus (three spined stickleback)
 ORGANISM Gasterosteus aculeatus (Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.
 1 (bases 1 to 1454)
 AUTHORS Kingsley, D., Grimwood, J., Dickson, M., Schmutz, J. and Myers, R.M.
 TITLE Expressed sequence tags from Gasterosteus aculeatus
 JOURNAL Unpublished (2004)
 COMMENT Contact: Grimwood, Jane
 Stanford Human Genome Center
 Stanford University School of Medicine
 975 S California Avenue, Palo Alto, CA 94304, USA
 Tel: 650 320 5917
 Fax: 650 320 5801
 Email: jane@hgsc.stanford.edu
 Plate: 13
 Class: BAC ends
 High quality sequence start: 30
 High quality sequence stop: 872.
 Location/Qualifiers
 1..1454
 /organism="Gasterosteus aculeatus"
 /mol_type="genomic DNA"
 /strain="Salmon River"
 /db_xref="taxon:69293"
 /clone="CH213-13H21"
 /sex="Mixed"
 /cell_type="Blood"
 /clone_1ib="CH213"
 /note="Vector: pTAR2.1; Site 1: EcoRI. The sequence of the clone was established as a mapping and sequencing collaboration at the Stanford Genome Evolution Center, funded by the NIH Centers of Excellence in Genomic Science (CEGS) initiative (<http://cegs.stanford.edu>). The clone was isolated from the BAC library CHOR-213 built by Peter deJong in collaboration with the Stanford Genome Evolution Center (<http://www.chori.org/bacpac/>). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm)."

ORIGIN

Query Match 1.8%; Score 23; DB 9; Length 1454;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 CATTGTCATCAGATGCATGCA 111
 ||||||||||||||||||||
 Db 961 CATTGTCATCAGATGCATGCA 983

RESULT 4
 LOCUS AQ663605 542 bp DNA 1linear GSS 23-JUN-1999
 DEFINITION HS_5476_A2_H04_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1052 Col=8 Row=O, genomic survey sequence.
 ACCESSION AQ663605
 VERSION AQ663605.1 GI:5171373
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 542)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, U., Young, U., Zhao, S., Adams, M.D. and Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: <http://www.htsc.washington.edu>
 Plate: 1052 Row: O Column: 8
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 542.
 Location/Qualifiers
 1..542
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=1052 Col=8 Row=O"
 /sex="male"
 /clone_1ib="RPCI-11 Human Male BAC Library"
 /note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

ORIGIN

Query Match 1.7%; Score 22; DB 8; Length 542;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 CATTCTCTGCGTGGATATTT 53
 ||||||||||||||||||||
 Db 160 CATTCTCTGCGTGGATATTT 139

RESULT 5
 LOCUS BM705592 331 bp mRNA 1linear EST 28-FEB-2002
 DEFINITION UI-E-DW0-agi-c-04-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
 UI-E-DW0-agi-c-04-0-UI 5', mRNA sequence.
 ACCESSION BM705592
 VERSION BM705592.1 GI:1901850
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 331)
 AUTHORS Bonaldi, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resegen.com).
 Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

```
1..331
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DW0-agi-c-04-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_1ib="UI-E-DW0"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-DW0 is a cDNA library containing the following
tissue(s): lens. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGATTGCGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

ORIGIN

Query Match 1.6%; Score 21; DB 4; Length 331;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTCTGCTG 1135
 |||||
 Db 55 GTGTGTTCCAGTCTGCTG 75

RESULT 6
 AV684301 349 bp mRNA linear EST 16-JAN-2002
 LOCUS AV684301 GKC Homo sapiens cDNA clone GKCA0B04 5', mRNA sequence.
 ACCESSION AV684301
 VERSION AV684301.1 GI:10286164
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 349)

TITLE

Insight into hepatocellular carcinogenesis at transcriptome level
 by comparing gene expression profiles of hepatocellular carcinoma
 with those of corresponding noncancerous liver
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

JOURNAL MEDLINE
 PUBMED 21625106
 COMMENT 11752456

COMMENT

Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922

FEATURES

source Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 Location/Qualifiers
 1..349
 /organism="Homo sapiens"

ORIGIN

Query Match 1.6%; Score 21; DB 1; Length 349;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTCTGCTG 1135
 |||||
 Db 180 GTGTGTTCCAGTCTGCTG 200

RESULT 7
 AV696353 349 bp mRNA linear EST 16-JAN-2002
 LOCUS AV696353 GKC Homo sapiens cDNA clone GKCAPB02 5', mRNA sequence.
 DEFINITION AV696353
 ACCESSION AV696353
 VERSION AV696353.1 GI:10298216
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 349)

TITLE

Insight into hepatocellular carcinogenesis at transcriptome level
 by comparing gene expression profiles of hepatocellular carcinoma
 with those of corresponding noncancerous liver
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

JOURNAL MEDLINE
 PUBMED 11752456
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922

FEATURES
 source Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 Location/Qualifiers
 1..349

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKCAPB02"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_1ib="GKC"
/notes="Vector: pBluescript sk(-); Site 1: EcoRI; Site 2:
XhoI"
```

ORIGIN

Query Match 1.6%; Score 21; DB 1; Length 349;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTCTGCTG 1135
 |||||
 Db 180 GTGTGTTCCAGTCTGCTG 200

RESULT 8
AV697455 349 bp mRNA 1linear EST 16-JAN-2002
LOCUS AV697455 GKC Homo sapiens cDNA clone GKACPC02 5', mRNA sequence.
DEFINITION AV697455
ACCESSION AV697455
VERSION AV697455.1 GI:10299318
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 349)
AUTHORS Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,
Shen, K., Lu, G., Pu, G., Zhong, W., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z. and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

JOURNAL MEDLINE
PUBMED
21625106
11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. 349
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKACPC02"
/cissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_1lb="GKC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 1.6%; Score 21; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1115 GTGTGTTCCAGTCTGTGCTG 1135
|||||
180 GTGTGTTCCAGTCTGTGCTG 200

DB 180 GTGTGTTCCAGTCTGTGCTG 200

RESULT 9
AV695717 360 bp mRNA 1linear EST 16-JAN-2002
LOCUS AV695717 GKC Homo sapiens cDNA clone GKACZG01 5', mRNA sequence.
DEFINITION AV695717
ACCESSION AV695717
VERSION AV695717.1 GI:10297580
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 360)
AUTHORS Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,
Shen, K., Lu, G., Pu, G., Zhong, W., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z. and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

MEDLINE
PUBMED
21625106
11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. 360
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKACZG01"
/cissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_1lb="GKC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 1.6%; Score 21; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1115 GTGTGTTCCAGTCTGTGCTG 1135
|||||
180 GTGTGTTCCAGTCTGTGCTG 200

DB 180 GTGTGTTCCAGTCTGTGCTG 200

RESULT 10
BI036922/c 368 bp mRNA 1linear EST 14-JUN-2001
LOCUS BI036922
DEFINITION MR4-NT0141-080101-209-e04 NT0141 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI036922
VERSION BI036922.1 GI:1443548
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 368)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE
PUBMED
10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR4&ct2=MR4-NT0141-
080101-209-e04&ct3=2001-01-08&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 35
High quality sequence stop: 345.
Location/Qualifiers
1. 368

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NT0141"
/note="Organ: nervous tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 1.6%; Score 21; DB 4; Length 368;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGTGTTCCAGGTCGTGTTG 1135
|||||
135 GTGTGTTCCAGGTCGTGTTG 115

RESULT 11 368 bp mRNA linear EST 04-MAR-2002
LOCUS BM762828
DEFINITION K-EST0044039 S12SNU216 Homo sapiens cDNA clone S12SNU216-12-C03 5',
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
ACCESSION BM762828
VERSION BM762828.1 GI:19092443
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 368)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 12 row: C column: 03
High quality sequence stop: 368.
Location/Qualifiers
1..368

FEATURES
source
1..368
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-12-C03"
/sex="F"
/tissue_type="lymph node"
/cell_type="Epithelial"
/cell_line="SNU-216"
/lab_host="T010F"
/clone_lib="S12SNU216"
/note="Organ: Stomach; Vector: pcns; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand

converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli T010F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 1.6%; Score 21; DB 4; Length 368;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGTGTTCCAGGTCGTGTTG 1135
|||||
209 GTGTGTTCCAGGTCGTGTTG 229

RESULT 12 369 bp mRNA linear EST 13-MAR-1995
LOCUS F12095
DEFINITION HSC35F071 normalized infant brain cDNA Homo sapiens cDNA clone
c-35f07, mRNA sequence.
ACCESSION F12095
VERSION F12095.1 GI:706428
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 369)
AUTHORS Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,
Devignes,M.D., Duprat,S., Houligatte,R., Jumeau,M.N., Lamy,B.,
Sebastien-Kabackchis,C. and Tessier,A.

TITLE IMAG: molecular integration of the analysis of the human genome
and its expression
JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
MEDLINE 95277534
PubMed 7757816
COMMENT Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 Evry Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_id: C; Genexpress_sequence_id: Y1C-35f07
Seq primer: (-21)M13 universal.
Location/Qualifiers
1..369

FEATURES
source
1..369
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c-35f07"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/clone_lib="normalized infant brain cDNA"
/note="Organ: brain; Vector: lafmd BA; Site 1: HindIII;
Site 2: NotI; sex=Female; dev stage=3 months old;
isolate=muscular atrophy patient; tissue type=total
brain; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmd BA vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

ORIGIN

Query Match 1.6%; Score 21; DB 7; Length 369;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGTGTTCCAGGTCGTGTTG 1135

Db 57 GTGGTTCACAGTCTGCTG 77

RESULT 13
LOCUS BP623179/c 371 bp mRNA linear EST 26-JUN-2004
DEFINITION BP623179 RAF117 Arabidopsis thaliana cDNA clone RAF117-05-N11 3', mRNA sequence.

ACCESSION BP623179 GI:49274361
VERSION BP623179.1
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 371)
Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Cono, Y., Muramatsu, M., Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shingawa, A. and Shinozaki, K.,
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)
21932900
PUBMED 11910074
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rcc.riken.go.jp
reversed clone; please visit our web site
(http://pigeonweb.gsc.riken.go.jp/) for further details.
Location/Qualifiers
1..371
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF117-05-N11"
/lab_host="DH10B"
/clone_1lb="RAF117"
/note="Site 1: BamHI; Site 2: SalI; Subtraction library.
The sequence was obtained from samples subjected to
dehydration-treated (1, 2, 5, 10 and 24 hr) and
rehydration-treated (1, 2, 5, 10, and 24 hr after
dehydration treatment)."

ORIGIN

Query Match 1.6%; Score 21; DB 5; Length 371;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1126 GGTGCTGATACGATGCA 1146
|||||
Db 280 GGTGCTGATACGATGCA 260

RESULT 14
LOCUS AW361182/c 372 bp mRNA linear EST 04-FEB-2000
DEFINITION RCI-CT0252-300999-011-c01 CT0252 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW361182
VERSION AW361182.1 GI:6865832
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 372)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project

JOURNAL
COMMENT Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC1&t2=RC1-CT0252-
300999-011-c01&t3=1999-09-30&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 44
High quality sequence stop: 245.
Location/Qualifiers
1..372
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/det_stage="Adult"
/clone_1lb="CT0252"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN

Query Match 1.6%; Score 21; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGGTTCACAGTCTGCTG 1135
|||||
Db 132 GTGGTTCACAGTCTGCTG 112

RESULT 15
LOCUS CF141588 386 bp mRNA linear EST 06-AUG-2003
DEFINITION UI-HP-CB0-aga-g-06-0-UI r1 NIH_MGC_210 Homo sapiens cDNA clone
IMAGE:3099803 5', mRNA sequence.
ACCESSION CF141588
VERSION CF141588.1 GI:33257032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 386)
Bonaldi, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
JOURNAL 8889548
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Tim Ratliff
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html

FEATURES Seq primer: PYX-5.
Location/Qualifiers
Source 1..386
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3099803"
/tissue_type="CNCAP(3)T-225 cell line"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH MGC 210"
/note="Organ: Prostate; Vector: pT73 Pac; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pT73 Pac vector. The library tag sequence located between the Not I site and the polyA tail is CCCAC. Tissue was provided by Tim Ratliff."

Query Match 1.6%; Score 21; DB 6; Length 386;
Best Local Similarity 100.0%; Pred. No. 29; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
Db 262 GTGTGTTCCAGTGTCTGTG 282

RESULT 16 CN345987 414 bp mRNA linear EST 16-MAY-2004
LOCUS CN345987
DEFINITION 1700600003542 GRN_PRENEM Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN345987
VERSION CN345987.1 GI:47345921
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 414)
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
CONTACT: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 414 Std Error: 0.00.
Location/Qualifiers
1..414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"
/clone_1lb="GRN_PRENEM"
/note="Oligo dT primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic acid and mitogens."

ORIGIN
Query Match 1.6%; Score 21; DB 7; Length 414;

Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
Db 359 GTGTGTTCCAGTGTCTGTG 379

RESULT 17 AA150852 424 bp mRNA linear EST 19-MAY-1997
LOCUS AA150852
DEFINITION 144b05.x1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone IMAGE:504753 5', similar to TR:G184462 G184462 CHAPERONIN-LIKE PROTEIN, ;, mRNA sequence.
ACCESSION AA150852
VERSION AA150852.1 GI:17222363
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 424)
Hillier, L., Lennon, G., Becker, M., Bernaldo, M.F., Chiappelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Maria, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 807 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham.
Location/Qualifiers
1..424
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3809741"
/db_xref="taxon:9606"
/clone="IMAGE:504753"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Soares pregnant uterus NBHPU"
/note="Organ: uterus; Vector: pT73-Pac; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', AACTGAGAAGATCGCGCGCCCTTTTCTTTTCTTTTCTTTT 3', double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by M. Fatima Bernaldo."

FEATURES

source

ORIGIN

Query Match 1.6%; Score 21; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1115 GTGTGTTCCAGTGTCTGTG 1135
|||||
Db 82 GTGTGTTCCAGTGTCTGTG 102

	RESULT 18
LOCUS	M69312 425 bp mRNA linear EST 16-OCT-1996
DEFINITION	zdd6g01.r1 Soares fetal heart NBHH19w Homo sapiens cDNA clone IMAGE:343728 5' similar to PIR:A46141 A46141 chaparonin-like protein HTR3 - human ;, mRNA sequence.
ACCESSION	M69312
VERSION	M69312.1 GI:1378573
KEYWORDS	EST.
SOURCE	Homo sapiens (human).
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 425)
AUTHORS	Hillier,L., Clark,N., Dubague,T., Elliston,K., Hawkins,M., Holtman,M., Holtman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,D., Rifkin,L., Ronfling,T., Soares,M., Tan,F., Trevaast,E., Waterston,R., Willamsen,A., Woldmann,P. and Wilson,R.
TITLE	The WashU-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK Washington Univeristy School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LINTL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert length: 1259 Std Error: 0.00 Seq primer: mob.RBGA+ET High quality sequence stop: High

[illegible]

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 426)
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,..., Naeai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P., Goldman,G.H., Carvalho,A.P., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	2002663
PIRMBED	10737800
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the PAPER/P/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&ct=CM3-HT0193-031199-030-f02&f3=1999-11-03&ct4=1) Seq primer: puc 18 forward High quality sequence start: 36 High quality sequence stop: 426. Location/Qualifiers
FEATURES	

[illegible]

TITLE Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 JOURNAL Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
 COMMENT Wilson, R.
 The WashU-Merck EST Project
 Unpublished (1995)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

FEATURES
 source
 High quality sequence stops: 362 Source: IMAGE Consortium, LNL This
 clone is available royalty-free through LNL; contact the IMAGE
 Consortium (info@image.llnl.gov) for further information.
 Insert length: 1405 Std Error: 0.00
 Seq primer: M13Rpi
 High quality sequence stop: 362.
 Location/Qualifiers
 1. 426
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:394259"
 /db_xref="taxon:9606"
 /clone="IMAGE:21912"
 /sex="female"
 /dev_stage="73 days post natal"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares infant brain LNB"
 /note="Organ: whole brain. Vector: lafield BA; Site 1: Not
 I; Site 2: Hind III; 1st strand cDNA was primed with a Not
 I - oligo(dT) primer [5']
 AACTGAGAGATTGGCGCCGACGAGATTTTTTTTTTTTTTTT 31;
 double-stranded cDNA was ligated to Hind III adaptors
 (Pharmacia), digested with Not I and directionally cloned
 into the Not I and Hind III sites of the lafield BA vector.
 Library went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 1.6%; Score 21; DB 7; Length 426;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1115 GTGTGTTCCAGTCTGTG 1135
 ||||||||||||||||||||
 DB 28 GTGTGTTCCAGTCTGTG 48

RESULT 21 434 bp mRNA linear EST 06-MAR-2002
 BM853177
 LOCUS K-EST0134548 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-134-C12
 DEFINITION 5', mRNA sequence.
 ACCESSION BM853177
 VERSION BM853177.1 GI:19209576
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 434)
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409

FEATURES
 source
 Email: yongsung@mail.kribb.re.kr
 Plate: 134 row: C column: 12
 High quality sequence stop: 434.
 Location/Qualifiers
 1. 434
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S22SNU16n1-134-C12"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /lab_host="SNU-16"
 /clone_lib="S22SNU16n1"
 /note="Organ: Stomach; Vector: pTZ19-Pac; Site 1: EcoRI;
 Site 2: NotI; The S22SNU16 library was contributed by
 Soares Laboratory and it was constructed as described by
 Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
 Research 6(9): 791-806. RNA was prepared from harvested
 cells of SNU-16 culture. SNU-16 cell was obtained from
 Korean Cell Line Bank (KCLB). SNU-16 was established from
 ascitic fluids of Korean patients by Park J.G. et al.
 (1990), Cancer Res 50: 2773-2780."

ORIGIN

Query Match 1.6%; Score 21; DB 4; Length 434;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1115 GTGTGTTCCAGTCTGTG 1135
 ||||||||||||||||||||
 DB 371 GTGTGTTCCAGTCTGTG 391

RESULT 22 434 bp mRNA linear EST 28-JAN-2003
 CB117296
 LOCUS K-EST0162463 L4SNU368 Homo sapiens cDNA clone L4SNU368-26-H04 5',
 DEFINITION mRNA sequence.
 ACCESSION CB117296
 VERSION CB117296.1 GI:27943103
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 434)
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 26 row: H column: 04
 High quality sequence stop: 434.
 Location/Qualifiers
 1. 434
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L4SNU368-26-H04"
 /sex="M"
 /tissue_type="Liver"
 /cell_type="Polygonal"
 /cell_line="SNU-368"
 /lab_host="Top10F"

/clone_11b="L4SNUS368"
 /note="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then deapped
 with tabacco acid pyrophosphatase (TAP). The deapped
 intact mRNA was ligated with DNA-RNA linker including
 EcoRI site by treatment of T4 RNA ligase and the first
 strand cDNA was synthesized from oligo dT-selected mRNA by
 priming with dT-tailed vector. The dT-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 434;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTCTGCTG 1135
 |||||
 Db 155 GTGTGTTCCAGTCTGCTG 175

RESULT 23
 A1910561 441 bp mRNA linear EST 28-JUL-1999
 LOCUS W68h07.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:2362525 3' similar to contains Alu repetitive
 element; contains element MER28 repetitive element; mRNA sequence.
 A1910561
 A1910561.1 GI:5630297

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 441)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 440.
 Location/Qualifiers

FEATURES
 source
 1..441
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2362525"
 /lab_host="DH10B"
 /clone_11b="Soares NFL T GBC S1"
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBH19, testis NHT, and B-cell
 NCI-CCAP GCB1) were mixed, and 88 circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 1.6%; Score 21; DB 1; Length 441;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1260 TGTACTGAGAAAAACCGA 1280
 |||||
 Db 335 TGTACTGAGAAAAACCGA 355

RESULT 24
 A1127935 443 bp mRNA linear EST 26-NOV-1996
 LOCUS 2114f07.r1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone
 DEFINITION IMAGE:501925 5' similar to TR:0187462 G184462 CHAPERONIN-LIKE
 PROTEIN; mRNA sequence.
 A1127935
 A1127935.1 GI:1687363

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 443)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevaekis, E., Waterston, R., Williamson, A., Wohlmann, P. and
 Wilson, R.
 The WashU-Merck EST Project
 Unpublished (1995)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watscn.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 311.
 Location/Qualifiers

FEATURES
 source
 1..443
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3806913"
 /db_xref="taxon:9606"
 /clone="IMAGE:501925"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_11b="Soares pregnant uterus NBHPU"
 /note="Organ: uterus; Vector: pT73D-Pac; Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer (5',
 AACTGAGAGATTTCGGCGCGCTTTTCTTTTCTTTTCTTTT 3',
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

ORIGIN

Query Match 1.6%; Score 21; DB 1; Length 443;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTCTGCTG 1135
 |||||
 Db 82 GTGTGTTCCAGTCTGCTG 102

RESULT 25

BG761673 459 bp mRNA linear EST 15-MAY-2001
 LOCUS 602718849P1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858669 5',
 DEFINITION mRNA sequence.
 BG761673
 ACCESSION BG761673.1 GI:14072326
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 459)
 NIH-MGC http://img.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/PTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1713 row: b column: 14
 High quality sequence stop: 457.
 Location/Qualifiers
 1..459
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4858669"
 /issue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="MD10B (phage-resistant)"
 /clone_lib="NIH_MGC_49"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 library."

ORIGIN
 Query Match 1.6%; Score 21; DB 4; Length 459;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGGTCTGTG 1135
 |||||||
 Db 107 GTGTGTTCCAGGTCTGTG 127

RESULT 26 493 bp mRNA linear EST 21-JUL-2000
 BE392663 601307517F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3625229 5',
 LOCUS mRNA sequence.
 BE392663
 ACCESSION BE392663.1 GI:9338028
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 493)
 NIH-MGC http://img.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM305 row: a column: 06
 High quality sequence stop: 493.
 Location/Qualifiers
 1..493
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3625229"
 /issue_type="endometrium, adenocarcinoma cell line"
 /lab_host="MD10B (phage-resistant)"
 /clone_lib="NIH_MGC_44"
 /note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
 Query Match 1.6%; Score 21; DB 2; Length 493;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGGTCTGTG 1135
 |||||||
 Db 330 GTGTGTTCCAGGTCTGTG 350

RESULT 27 499 bp mRNA linear EST 06-MAR-2002
 BM641588 K-EST0118850 S12SN216 Homo sapiens cDNA clone S12SN216-53-B11 5',
 LOCUS mRNA sequence.
 BM641588
 ACCESSION BM641588.1 GI:19197997
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 499)
 KIM,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.W., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongseung@mail.kribb.re.kr
 Plate: 53 row: B column: 1199.
 High quality sequence stop: 499.
 Location/Qualifiers
 1..499
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S12SN216-53-B11"
 /sex="F"
 /issue_type="lymph node"
 /tissue_type="Epithelial"
 /cell_line="SNU-216"

/lab_host="Top10F"
/clone_1ib="S12SN216"
/note="Organ: Stomach; Vector: pcn5; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 1.6%; Score 21; DB 4; Length 499;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTCTGTG 1135
|||||
Db 209 GTGTGTTCCAGTCTGTG 229

RESULT 28
BE265418 504 bp mRNA linear EST 13-JUL-2000
LOCUS 601194958P1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:353856 5',
DEFINITION mRNA sequence.
ACCESSION BE265418
VERSION BE265418.1 GI:9138986
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 504)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Plate: L1CM224 row: d column: 15
High quality sequence stop: 504.
Location/Qualifiers
1..504

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3538526"
/issue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_7"
/note="Organ: lung; Vector: pOTR7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAACAG(G). Size-selected >800bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 1.6%; Score 21; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTCTGTG 1135
|||||
Db 67 GTGTGTTCCAGTCTGTG 87

RESULT 29
CN345984 506 bp mRNA linear EST 16-MAY-2004
LOCUS 17000600169383 GRN_PRENEM Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN345984
ACCESSION CN345984
VERSION CN345984.1 GI:47345918
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 506)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Bradenberger, R., Wei, H., Zhang, S., Lei, S., Murae, J., Fisk, G.J.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Bradenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbradenberger@geron.com
Insert length: 506 Std Error: 0.00.
Location/Qualifiers
1..506

TITLE

JOURNAL

COMMENT

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_1ib="GRN_PRENEM"
/note="Oligo dt-primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."

ORIGIN

Query Match 1.6%; Score 21; DB 7; Length 506;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTCTGTG 1135
|||||
Db 466 GTGTGTTCCAGTCTGTG 506

RESULT 30
CB159634 511 bp mRNA linear EST 30-JAN-2003

LOCUS K-EST0219212 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-25-B04
DEFINITION 5', mRNA sequence.
ACCESSION CB159634
VERSION CB159634.1 GI:28145760
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 511)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Bukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.
ZIC Frontiers Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 25 row: B column: 04
High quality sequence stop: 511.
Location/Qualifiers
1. 511

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L18POOL1n1-25-B04"
/cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
/lab_host="Top10F"
/clone_lib="L18POOL1n1"
/note="Organ: Liver; Vector: pTTT3-Pac; Site_1: EcoRI;
Site_2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN
Query Match 1.6%; Score 21; DB 6; Length 511;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTGCTGTG 1135
|||||
Db 443 GTGTGTTCCAGTGTGCTGTG 463

RESULT 31
BM796788 516 bp mRNA linear EST 05-MAR-2002
LOCUS K-EST0079591 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-63-H11
DEFINITION 5', mRNA sequence.
ACCESSION BM796788
VERSION BM796788.1 GI:19145020
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 516)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.U., Cheong,U.E., Sohn,H.Y., Kim,J.W., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS

TITLE
JOURNAL
COMMENT
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 63 row: H column: 11
High quality sequence stop: 516.
Location/Qualifiers
1. 516

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S22SNU16n1-63-H11"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-16"
/lab_host="DH10B"
/clone_lib="S22SNU16n1"

/note="Organ: Stomach; Vector: pTTT3-Pac; Site_1: EcoRI;
Site_2: NotI; The S22SNU16 library was contributed by the
Soares laboratory and it was constructed as described by
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. RNA was prepared from harvested
cells of SNU-16 culture. SNU-16 cell was obtained from
Korean Cell Line Bank (KCLB). SNU-16 was established from
ascitic fluids of Korean patients by Park J.G. et al.
(1990), Cancer Res 50: 2773-2780."

ORIGIN
Query Match 1.6%; Score 21; DB 4; Length 516;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTGCTGTG 1135
|||||
Db 321 GTGTGTTCCAGTGTGCTGTG 341

RESULT 32
BG149625 522 bp mRNA linear EST 05-FEB-2001
LOCUS nad31d12.x1 NCI CGAP Lu24 Homo sapiens cDNA IMAGE:336983 3'
DEFINITION similar to SW:TCBPZ_HUMAN P40227 T-COMPLEX PROTEIN 1, ZETA SUBUNIT
; , mRNA sequence.
ACCESSION BG149625
VERSION BG149625.1 GI:12661655
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 522)
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbcr@mail.nih.gov
Tissue Procurement: Christopher Moshaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL, send email to:
info@image.lnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 460.
Location/Qualifiers
1. 522

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:336983"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI-CGAP Lu24"
/note="Organ: lung; Vector: pTTT3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP Lu5 was prepared, and ss circles was
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneids
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo.

ORIGIN
Query Match 1.6%; Score 21; DB 4; Length 522;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTCTGGTG 1135
 |||||
 DB 458 GTGTGTTCCAGTGTCTGGTG 478

RESULT 33
 CB124236 528 bp mRNA linear EST 29-JAN-2003
 LOCUS K-EST0172905 L5HLK1b1 Homo sapiens cDNA clone L5HLK1b1-21-B04 5',
 DEFINITION mRNA sequence.
 CB124236
 VERSION CB124236.1 GI:28084249
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 528)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Kim,Y.S., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Oh,K.J.

TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 21 row: B column: 04
 High quality sequence stop: 528.

FEATURES

source

1..528
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L5HLK1b1-21-B04"
 /sex="M"
 /cell_line="HLK-1"
 /lab_host="Top10P"
 /clone_1ib="L5HLK1b1"
 /note="Organ: Liver; Vector: PCNS-D2; Site 1: EcoRI;
 Site 2: NotI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then deacapped
 with tobacco acid pyrophosphatase (TAP). The deacapped
 intact mRNA was ligated with DNA-RNA linker including
 EcoRI site by treatment of T4 RNA ligase and the first
 strand cDNA was synthesized from oligo dT-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10P' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library. After analyzing and
 sequencing about 2,000 - 3,000 colonies in original cDNA
 library, the abundant cDNAs were selected and amplified by
 PCR reaction using vector region primer including T7
 promoter as 5' primer and NidFT14 as 3' primer. The PCR
 products were used as template for synthesis of
 biotinylated single stranded RNA by in vitro transcription
 reaction. The synthesized RNA probes were hybridized with
 antisense single stranded cDNAs prepared from original
 library and incubated with avidin-gel. After removing
 DNA-RNA hybriide by centrifuge, the substracted cDNA
 libraries were constructed by transformation of the
 remaining DNA into competent cells E. coli Top10P' with
 electroporation method."

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 528;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTCTGGTG 1135
 |||||
 DB 343 GTGTGTTCCAGTGTCTGGTG 363

RESULT 34
 W38880 536 bp mRNA linear EST 15-MAY-1996
 LOCUS zb32a05.r1 Soares parathyroid tumor NbHPa Homo sapiens cDNA clone
 DEFINITION IMAGE:305264 5' similar to PIR:A46141 A46141 chaperonin-1like
 protein HHR3 - human ;, mRNA sequence.
 W38880
 VERSION W38880.1 GI:1320586
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 536)
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawking,M.,
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
 Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
 Trevaaskis,E., Waterston,R., Willmsom,A., Wohlmann,P. and
 Wilson,R.

TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: mob.REGA+ET
 High quality sequence stop: 383.

FEATURES

source

1..536
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1248674"
 /db_xref="taxon:9606"
 /clone="IMAGE:305264"
 /tissue_type="parathyroid tumor"
 /dev_stage="adult"
 /lab_host="PH10B (ampicillin resistant)"
 /clone_1ib="Soares parathyroid tumor NbHPa"
 /note="Organ: parathyroid gland; Vector: pT73D
 (pharmacia) with a modified polylinker; Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer
 (5'-GTGTACAACTCTGAAGTGGAGCGGCCGACACATTTTCTTTTCTTTT
 TTTT-3'), double-stranded cDNA was size selected, ligated
 to Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of a modified pT73
 vector (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
 adenomas was kindly provided by Dr. Stephen Marx, National
 Institute of Diabetes and Digestive and Kidney Diseases,
 NIH."

ORIGIN

Query Match 1.6%; Score 21; DB 7; Length 536;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1115 GTGTGTTCCAGTGTCTGGTG 1135

Db 79 GTGTGTTCCAGGTCTGTGTG 99

RESULT 35
BM840031

LOCUS

DEFINITION BM840031 542 bp mRNA linear EST 06-MAR-2002
K-EST0117031 S12SNU216 Homo sapiens cDNA clone S12SNU216-31-B03 5',
mRNA sequence.

ACCESSION BM840031

VERSION BM840031.1 GI:19196440

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 542)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Beon-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 31 row: B column: 03
High quality sequence stop: 542.

FEATURES
source
1..542
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-31-B03"
/sex="F"
/issue_type="lymph node"
/cell_line="Epithelial"
/cell_line="SNU-216"
/lab_host="Top10F"
/lab_host="Top10F"
/clone_lib="S12SNU216"
/note="Organ: Stomach; Vector: pCNS; Site:1: EcoRI;
Site_2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The deacapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Beig method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 1.6%; Score 21; DB 4; Length 542;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1115 GTGTGTTCCAGGTCTGTGTG 1135
|||||
|||||
209 GTGTGTTCCAGGTCTGTGTG 229

Db

RESULT 36
CB119331

LOCUS

CB119331 543 bp mRNA linear EST 28-JAN-2003

DEFINITION K-EST0166075 L10choCK0 Homo sapiens cDNA clone L10choCK0-9-B04 5',
mRNA sequence.

ACCESSION CB119331

VERSION CB119331.1 GI:27945135

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 543)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Beon-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 9 row: E column: 04
High quality sequence stop: 543.

FEATURES
source
1..543
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L10choCK0-9-B04"
/sex="M"
/cell_line="Cho-CK"
/lab_host="Top10F"
/clone_lib="L10choCK0"
/note="Organ: Liver; Vector: pT73-Pac; Site:1: EcoRI;
Site_2: NotI; The library was contributed by the Soares
Laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 543;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1115 GTGTGTTCCAGGTCTGTGTG 1135
|||||
|||||
505 GTGTGTTCCAGGTCTGTGTG 525

Db

RESULT 37
CB147674

LOCUS

DEFINITION CB147674 547 bp mRNA linear EST 29-JAN-2003
K-EST0203727 L14choiCK0 Homo sapiens cDNA clone L14choiCK0-17-G06
5', mRNA sequence.

ACCESSION CB147674

VERSION CB147674.1 GI:28129145

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 547)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Beon-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 17 row: G column: 06
 High quality sequence stop: 547.
 Location/Qualifiers

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Umbilical vein endothelial cells, passaged once. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

FEATURES
 source
 1..547
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="U14choiCK0-17-G06"
 /sex="M"
 /cell_line="Choi-CK"
 /lab_host="Top10F"
 /clone_lib="U14choiCK0"
 /note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 547;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTCTGTGTG 1135
 |||||
 Db 388 GTGTGTTCCAGTCTGTGTG 408

RESULT 38
 AA186456 548 bp mRNA linear EST 10-MAR-1998
 LOCUS zpt0h10.r1 Stratsgene endothelial cell 937223 Homo sapiens cDNA
 DEFINITION clone IMAGE:625603 5' similar to TR:G184462 G184462 CHAPERONIN-LIKE
 PROTEIN, mRNA sequence.
 ACCESSION AA186456
 VERSION AA186456.1 GI:1774556
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 548)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,
 Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
 Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
 Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: esewatson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 931 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amerham
 High quality sequence stop: 301.
 Location/Qualifiers

FEATURES

source
 1..548
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:5047757"
 /db_xref="taxon:9606"
 /clone="IMAGE:625603"
 /dev_stage="umbilical vein, 1 passage"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratsgene endothelial cell 937223"

ORIGIN
 Query Match 1.6%; Score 21; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTCTGTGTG 1135
 |||||
 Db 112 GTGTGTTCCAGTCTGTGTG 132

RESULT 39
 CB160975 553 bp mRNA linear EST 30-JAN-2003
 LOCUS K-EST0220750 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-24-B06
 DEFINITION 5' mRNA sequence.
 ACCESSION CB160975
 VERSION CB160975.1 GI:28147101
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 553)
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 24 row: B column: 06
 High quality sequence stop: 553.
 Location/Qualifiers

FEATURES

source
 1..553
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L18POOL1n1-24-B06"
 /cell_line="SNU-354+choi-CK+HLK-3"
 /lab_host="Top10F"
 /clone_lib="L18POOL1n1"
 /note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

ORIGIN

Query Match 1.6%; Score 21; DB 6; Length 553;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1115 GTGTGTTCCAGTCTGTGTG 1135
 |||||
 Db 427 GTGTGTTCCAGTCTGTGTG 447

RESULT 40
 AO837961/c 553 bp DNA linear GSS 30-AUG-1999
 LOCUS AO837961


```
977 15 1.2 3585 1 US-08-455-543A-34 Sequence 34, Appl
978 15 1.2 3585 2 US-08-193-078B-23 Sequence 23, Appl
979 15 1.2 3585 2 US-08-223-305C-34 Sequence 34, Appl
980 15 1.2 3585 3 US-08-949-386-30 Sequence 30, Appl
981 15 1.2 3585 3 US-08-450-562-30 Sequence 30, Appl
982 15 1.2 3585 3 US-08-984-709A-30 Sequence 30, Appl
983 15 1.2 3585 3 US-08-450-272-30 Sequence 30, Appl
984 15 1.2 3585 4 US-08-450-272-30 Sequence 30, Appl
985 15 1.2 3585 4 US-08-450-272-30 Sequence 30, Appl
986 15 1.2 3585 4 US-08-450-272-30 Sequence 30, Appl
987 15 1.2 3600 1 US-08-455-543A-11 Sequence 11, Appl
988 15 1.2 3600 1 US-08-193-078B-11 Sequence 11, Appl
989 15 1.2 3600 2 US-08-223-305C-11 Sequence 11, Appl
990 15 1.2 3600 2 US-08-223-305C-11 Sequence 11, Appl
991 15 1.2 3600 2 US-08-149-097D-11 Sequence 11, Appl
992 15 1.2 3600 3 US-08-949-386-11 Sequence 11, Appl
993 15 1.2 3600 3 US-08-450-562-11 Sequence 11, Appl
994 15 1.2 3600 3 US-08-984-709A-11 Sequence 11, Appl
995 15 1.2 3600 3 US-08-450-272-11 Sequence 11, Appl
996 15 1.2 3600 4 US-08-450-272-11 Sequence 11, Appl
997 15 1.2 3601 3 US-09-017-631-23 Sequence 23, Appl
998 15 1.2 3602 2 US-08-883-795A-33 Sequence 33, Appl
999 15 1.2 3602 3 US-09-018-138-1 Sequence 1, Appl
1000 15 1.2 3628 3 US-08-480-640A-113 Sequence 113, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-792-024-42
Sequence 42, Application US/09792024
Patent No. 6783985
GENERAL INFORMATION:
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Gene Disruption Methodologies for Drug
FILE REFERENCE: 10182-004-999
CURRENT APPLICATION NUMBER: US/09/792,024
CURRENT FILING DATE: 2001-02-20
NUMBER OF SEQ ID NOS: 490
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42
LENGTH: 1299
TYPE: DNA
ORGANISM: Candida albicans
US-09-792-024-42
```

Query Match 68.6%; Score 891; DB 4; Length 1299;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 1 ATGTCAAAAGCATTTAGTGCACCTGGAAGGCAATTTCTGCTGGTGATATTGGTTCTT 60
DB 1 ATGTCAAAAGCATTTAGTGCACCTGGAAGGCAATTTCTGCTGGTGATATTGGTTCTT 60
QY 61 GAGCCAAATTTAGTGCATTTAGTGCACGATTTGCTGCTGGTGATATTGGTTCTT 60
DB 61 GAGCCAAATTTAGTGCATTTAGTGCACGATTTGCTGCTGGTGATATTGGTTCTT 60
QY 121 CCAAAAGGACCAAGTTGAAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTTGCAAC 180
DB 121 CCAAAAGGACCAAGTTGAAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTTGCAAC 180
QY 181 GGAAGATGGGAATATCATATCATCAATACAGAGAGCCAGAGAGTTCACTACGCG 240
DB 181 GGAAGATGGGAATATCATATCATCAATACAGAGAGCCAGAGAGTTCACTACGCG 240
QY 241 ATAAATCCATTTTGAAGGCAACTATATTCATCGTTTATTTAGTTATTTCAACGACGAA 300
DB 241 ATAAATCCATTTTGAAGGCAACTATATTCATCGTTTATTTAGTTATTTCAACGACGAA 300
```

```
DB 241 ATAAATCCATTTTGAAGGCAACTATATTCATCGTTTATTTAGTTATTTCAACGACGAA 300
QY 301 GCATTTGATCTTGAATATCATATTTACTCAGACCTCGATATTCATTCAGAGAAATCT 360
DB 301 GCATTTGATCTTGAATATCATATTTACTCAGACCTCGATATTCATTCAGAGAAATCT 360
QY 361 GAAACCAAGCATCTCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GAAACCAAGCATCTCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 GAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 GAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 TTAATATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGGCAACAGT 540
DB 481 TTAATATCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGGCAACAGT 540
QY 541 GCAAGATTTGCACTGTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 GCAAGATTTGCACTGTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GCAATTTATGCTGATTTGATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 GCAATTTATGCTGATTTGATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 CAGGTTCTAGAAAGTGAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 CAGGTTCTAGAAAGTGAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 AACTGGGAATTCAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 AACTGGGAATTCAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 GAGCTCAAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 781 GAGCTCAAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 GAAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 GAAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 ATGAAGGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 ATGAAGGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 GAGTTAGATCATTTCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 961 GAGTTAGATCATTTCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 CAAGCATTTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 CAAGCATTTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 GACCGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 GACCGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 GATGCAATGCTGTTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 GATGCAATGCTGTTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GAAATTCAGATTTATTTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 GAAATTCAGATTTATTTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 GTAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1299
DB 1261 GTAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1299
```

RESULT 2
US-09-248-796A-3819

```
; Sequence 3819, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3819
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1194)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-248-796A-3819
```

```
Query Match      44.1%; Score 573; DB 4; Length 1230;
Best Local Similarity 100.0%; Pred. No. 4.8e-276;
Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 586 TTTGATGTTGCACTGCAATTTATGTCGATTTGTTATATGAAGATTTTACGCCAGCTTTG 645
DB 613 TTTGATGTTGCACTGCAATTTATGTCGATTTGTTATATGAAGATTTTACGCCAGCTTTG 672
QY 646 ATAAATGACGTTTCAAGTTCTAGAAAGTGAATCCTGAGAAAGTTCCCAAGAGTTGAA 705
DB 673 ATAAATGACGTTTCAAGTTCTAGAAAGTGAATCCTGAGAAAGTTCCCAAGAGTTGAA 732
QY 706 AAATGATTTGAAGTAACTGGGAATTCAGAAATGAAGATGATACATTAACATACGGAATC 765
DB 723 AAATGATTTGAAGTAACTGGGAATTCAGAAATGAAGATGATACATTAACATACGGAATC 792
QY 766 AAGTTATTTAATGGGTGACGTCAGAGGTGCTCAGAAACCCCAATTGGTATCAAGATTA 825
DB 793 AAGTTATTTAATGGGTGACGTCAGAGGTGCTCAGAAACCCCAATTGGTATCAAGATTA 852
QY 826 CTCGAATGGAAGAAAGGAAAGCCAGAAAGAGCTCTGTTGATGACCGAGTTAATAGT 885
DB 853 CTCGAATGGAAGAAAGGAAAGCCAGAAAGAGCTCTGTTGATGACCGAGTTAATAGT 912
QY 886 GCCAATTTACAGTTTATGAAAGAAATGAGGAAATGCGTGAATAATACGACTCAGACCCA 945
DB 913 GCCAATTTACAGTTTATGAAAGAAATGAGGAAATGCGTGAATAATACGACTCAGACCCA 972
QY 946 GAGACTTATATTAAGAGTTAATCATCTTCTGTTGAGCCTTTGACTGTTGCGATTAAAGAC 1005
DB 973 GAGACTTATATTAAGAGTTAATCATCTTCTGTTGAGCCTTTGACTGTTGCGATTAAAGAC 1032
QY 1006 ATCAGAAAAGGTTTCAAGCATTTACACAAAATCAGAGGTTCCATTAAGACCTGATGTC 1065
DB 1033 ATCAGAAAAGGTTTCAAGCATTTACACAAAATCAGAGGTTCCATTAAGACCTGATGTC 1092
QY 1066 CAACCCAGTTTGGACCGTTGTCAAGAGATTCCTGTTGTTGTTGTTGTTGTTGTTGTTG 1125
DB 1093 CAACCCAGTTTGGACCGTTGTCAAGAGATTCCTGTTGTTGTTGTTGTTGTTGTTGTTG 1152
QY 1126 GGTGCTGGTGGATGACGATGCAATAGCTGTAATTA 1158
DB 1153 GGTGCTGGTGGATGACGATGCAATAGCTGTAATTA 1185
```

```
RESULT 3
US-09-248-796A-3820
; Sequence 3820, Application US/09248796A
; Patent No. 6747137
```

```
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3820
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (762)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-248-796A-3820
```

```
Query Match      28.2%; Score 366; DB 4; Length 765;
Best Local Similarity 99.1%; Pred. No. 9.5e-173;
Matches 666; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 ATGTCAAAAAGCATTTAGTGACCTGGAAGAAAGCATTTCTGCTGGTGATATTTGGTTCTT 60
DB 64 ATGTCAAAAAGCATTTAGTGACCTGGAAGAAAGCATTTCTGCTGGTGATATTTGGTTCTT 123
QY 61 GAGCAATTTATGATGCTTATGTTGACAGATTTGTCATCAGAAATGATGATTTATACA 120
DB 124 GAGCAATTTATGATGCTTATGTTGACAGATTTGTCATCAGAAATGATGATTTATACA 183
QY 121 CCAAAAAGAACCGATTTGAAAGAAATCTAGAAATCAAAATTTCTTCAACCCCAATTTGCAAC 180
DB 184 CCAAAAAGAACCGATTTGAAAGAAATCTAGAAATCAAAATTTCTTCAACCCCAATTTGCAAC 243
QY 181 GGAGATGGGAATATACATATCATCAATCAATACAGAAAGCCAGAGAGTTGATGATCAGC 240
DB 244 GGAGATGGGAATATACATATCATCAATCAATACAGAAAGCCAGAGAGTTGATGATCAGC 303
QY 241 ATAAATTCATTTTATGAGGCACTATATTCATCTGTTTATGATTAATTTCAACCGAAC 300
DB 304 ATAAATTCATTTTATGAGGCACTATATTCATCTGTTTATGATTAATTTCAACCGAAC 363
QY 301 GCATTTGATCTTGAATCATCTATTTACTCAGACCTCGATATCATTTCAAGAAAGTACT 360
DB 364 GCATTTGATCTTGAATCATCTATTTACTCAGACCTCGATATCATTTCAAGAAAGTACT 423
QY 361 GAAACCAAGACATCTCGAATGAGAAAGAAACATTTCTTATCATTTCTGTCGCAATTACC 420
DB 424 GAAACCAAGACATCTCGAATGAGAAAGAAACATTTCTTATCATTTCTGTCGCAATTACC 483
QY 421 GAAGTGAAGAAAGCCGATTAAGTTTATCGGACGAGATTAAGTTCAGTTGTTGCCAAGT 480
DB 484 GAAGTGAAGAAAGCCGATTAAGTTTATCGGACGAGATTAAGTTCAGTTGTTGCCAAGT 543
QY 481 TTAATTAATCCATTTATCCCAATGTTATCAGTACGAATAAAGATATTTTGCACACGTT 540
DB 544 TTAATTAATCCATTTATCCCAATGTTATCAGTACGAATAAAGATATTTTGCACACGTT 603
QY 541 GCACAGATTTGACATTTGTTATGCCCCAAAAGAAAGATGAGATCGGGTTGATGTTGCAACT 600
DB 604 GCACAGATTTGACATTTGTTATGCCCCAAAAGAAAGATGAGATCGGGTTGATGTTGCAACT 663
QY 601 GCAATTTATGCTGATTTGATATAGAGATTTTCAAGCAGCTTTGATTAATACGTTGTT 660
DB 664 GCAATTTATGCTGATTTGATATAGAGATTTTCAAGCAGCTTTGATTAATACGTTGTT 723
QY 661 CAGGTTCTAGAA 672
DB 724 CAGGTTCTAGAA 735
```

RESULT 4

US-09-792-024-348/c
; Sequence 348, Application US/09792024
; Patent No. 6783985
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jlang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug
; FILE REFERENCE: 10182-004-999
; CURRENT APPLICATION NUMBER: US/09/792,024
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 490
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 90
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA primer
US-09-792-024-348

Query Match

Best Local Similarity 5.0%; Score 65; DB 4; Length 90;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCAAGCATTTAGTGCACCGGAAAGCATTTCTGCTGCTGATTTGTTCTT 60
DB 65 ATGTCAAGCATTTAGTGCACCGGAAAGCATTTCTGCTGCTGATTTGTTCTT 6

QY 61 GAGCC 65
DB 5 GAGCC 1

RESULT 5

US-09-792-024-470/c
; Sequence 470, Application US/09792024
; Patent No. 6783985
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jlang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug
; FILE REFERENCE: 10182-004-999
; CURRENT APPLICATION NUMBER: US/09/792,024
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 490
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 470
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA primer
US-09-792-024-470

Query Match

Best Local Similarity 1.7%; Score 22; DB 4; Length 22;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 GGATTAGTGTACAGTTGTTGCCA 475
DB 22 GGATTAGTGTACAGTTGTTGCCA 1

RESULT 6

US-09-910-009A-157

Sequence 157, Application US/09910009A
Patent No. 6727226
GENERAL INFORMATION:
APPLICANT: University of Utah Research Foundation
APPLICANT: Cognetix, Inc.
APPLICANT: Oliveira, Baldomero M.
APPLICANT: McIntosh, J. Michael
APPLICANT: Garrett, James E.
APPLICANT: Watkins, Maren
APPLICANT: Cruz, Lourdes J.
APPLICANT: Shon, Ki-Joon
APPLICANT: Jacobsen, Richard
APPLICANT: Jones, Robert M.
APPLICANT: Carlier, G. Edward
APPLICANT: Shen, Greg S.
APPLICANT: Wasgstaff, John D.
TITLE OF INVENTION: Mg-Conopeptides
FILE REFERENCE: 2314-242
CURRENT APPLICATION NUMBER: US/09/910,009A
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 60/219,619
PRIOR FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: US 60/245,157
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/264,319
PRIOR FILING DATE: 2001-01-29
PRIOR APPLICATION NUMBER: US 60/277,270
PRIOR FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 520
SOFTWARE: PatentIn version 3.0
SEQ ID NO 157
LENGTH: 355
TYPE: DNA
ORGANISM: *Conus marmoreus*
US-09-910-009A-157

Query Match

Best Local Similarity 1.5%; Score 19; DB 4; Length 355;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 861 TGTTGTGTATGACCACTT 879
DB 262 TGTTGTGTATGACCACTT 280

RESULT 7

US-09-248-796A-2118/c
Sequence 2118, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 2118
LENGTH: 558
TYPE: DNA
ORGANISM: *Candida albicans*
US-09-248-796A-2118

Query Match

Best Local Similarity 1.5%; Score 19; DB 4; Length 558;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1091 AAGAGATTCCTGCTTGTCT 1109

Db 42 AAGAGATTCCTGTGTGT 24

RESULT 8

US-09-495-050A-89/C
; Sequence 89, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED F
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 89
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 1330882CT1
US-09-495-050A-89

Query Match 1.5%; Score 19; DB 4; Length 593;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 GAGAAAAACATTTCTTTA 401
Db 432 GAGAAAAACATTTCTTTA 414

RESULT 9

US-08-803-973-3/C
; Sequence 3, Application US/08803973
; Patent No. 5773701
; GENERAL INFORMATION:
; APPLICANT: Braun, Carl J.
; APPLICANT: Hemenway, Cynthia L.
; APPLICANT: Turner, Nilgun E.
; TITLE OF INVENTION: Plants Resistant to Infection by PVX
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/803,973
; FILING DATE: 21-FEB-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/804,862
; FILING DATE: 04-DEC-1991
; APPLICATION NUMBER: US 07/771,912
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10541)A

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-803-973-3

Query Match 1.4%; Score 18; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1101 TGGTGTGTGTGTGTGT 1118
Db 32 TGGTGTGTGTGTGTGT 15

RESULT 10

US-08-803-972-3/C
; Sequence 3, Application US/08803972
; Patent No. 5792937
; GENERAL INFORMATION:
; APPLICANT: Braun, Carl J.
; APPLICANT: Hemenway, Cynthia L.
; APPLICANT: Turner, Nilgun E.
; TITLE OF INVENTION: Plants Resistant to Infection by PVX
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/803,972
; FILING DATE: 21-FEB-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/804,862
; FILING DATE: 04-DEC-1991
; APPLICATION NUMBER: US 07/771,912
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10541)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-803-972-3

Query Match 1.4%; Score 18; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1101 TGGTGTGTGTGTGTGT 1118

Db 32 TGTGTGTGTGTGTGTGT 15

|||||

RESULT 11
US-08-956-171E-1674/c
; Sequence 1674, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunach
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 1674:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1674:
US-08-956-171E-1674

Query Match 1.4%; Score 18; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1105 TGTGTGTGTGTGTGT 1122
|||||

Db 66 TGTGTGTGTGTGTGT 49

RESULT 12
US-08-781-986A-1674/c
; Sequence 1674, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunach
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 1674:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-1674

Query Match 1.4%; Score 18; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1105 TGTGTGTGTGTGTGT 1122
|||||

Db 66 TGTGTGTGTGTGTGT 49

RESULT 13
US-08-956-171E-4078
; Sequence 4078, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunach
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
TELECOMMUNICATION INFORMATION:
REFERENCE/DOCKET NUMBER: PB248P1
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 4078:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4078:
US-08-956-171E-4078

Query Match 1.4%; Score 18; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 948 GACTTATATTAAGATT 965
Db 62 GACTTATATTAAGATT 79

RESULT 14
US-08-781-986A-4078
Sequence 4078, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunach
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 4078:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-4078

Query Match 1.4%; Score 18; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 948 GACTTATATTAAGATT 965
Db 62 GACTTATATTAAGATT 79

RESULT 15
US-08-956-171E-3865/C
Sequence 3865, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunach
Gail H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 3865:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3865:
US-08-956-171E-3865

Query Match 1.4%; Score 18; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 948 GACTTATATTAAGATT 965
Db 145 GACTTATATTAAGATT 128

RESULT 16
US-08-956-171E-3933
Sequence 3933, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunach
Gail H. Choi
Patrick S. Dillon
Craig A. Rosen

```

;
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 3933:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3933:
;
; US-08-956-171E-3933
;
; Query Match 1.4%; Score 18; DB 4; Length 400;
; Best Local Similarity 100.0%; Pred. No. 43;
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 948 GACTTATATTAAAGATT 965
; DB 18 GACTTATATTAAAGATT 35
;
; RESULT 17
; US-08-956-171E-4018
; Sequence 4018, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33

```

```

;
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 4018:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 4018:
;
; US-08-956-171E-4018
;
; Query Match 1.4%; Score 18; DB 4; Length 400;
; Best Local Similarity 100.0%; Pred. No. 43;
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 948 GACTTATATTAAAGATT 965
; DB 266 GACTTATATTAAAGATT 283
;
; RESULT 18
; US-08-781-986A-3865/c
; Sequence 3865, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8512
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 3865:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-3865

Query Match 1.4%; Score 18; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 GACTTATATTAAAGATT 965
DB 145 GACTTATATTAAAGATT 128

RESULT 19
US-08-781-986A-3933
Sequence 3933; Application US/08781986A
Patent No. 6737248

GENERAL INFORMATION:
APPLICANT: Charles Kunesh
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 3933:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-3933

Query Match 1.4%; Score 18; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 GACTTATATTAAAGATT 965
DB 18 GACTTATATTAAAGATT 35

RESULT 20
US-08-781-986A-4018
Sequence 4018; Application US/08781986A
Patent No. 6737248

GENERAL INFORMATION:
APPLICANT: Charles Kunesh
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 4018:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-4018

Query Match 1.4%; Score 18; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 GACTTATATTAAAGATT 965
DB 266 GACTTATATTAAAGATT 283

RESULT 21
US-09-543-681A-3991
Sequence 3991; Application US/09543681A
Patent No. 6605709

GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 3991
LENGTH: 1431
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-3991

Query Match 1.4%; Score 18; DB 4; Length 1431;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1241 TCGAAGACCAACAGAG 1258
DB 20 TCGAAGACCAACAGAG 37

RESULT 22

US-09-809-665A-17/c
; Sequence 17, Application US/09809665A
; Patent No. 6790950
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28341/00435
; CURRENT APPLICATION NUMBER: US/09/809,665A
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 3247
; TYPE: DNA
; ORGANISM: Pasteurella multocida
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1479)..(3245)
; OTHER INFORMATION: fhnb2
US-09-809-665A-17

Query Match
Best Local Similarity 1.4%; Score 18; DB 4; Length 3247;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 669 AGAAGTATCTCTGAGAA 686
Db 1443 AGAAGTATCTCTGAGAA 1426

RESULT 23
US-09-809-665A-19/c
; Sequence 19, Application US/09809665A
; Patent No. 6790950
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28341/00435
; CURRENT APPLICATION NUMBER: US/09/809,665A
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 3247
; TYPE: DNA
; ORGANISM: Pasteurella multocida
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1446)
; OTHER INFORMATION: fhac
US-09-809-665A-19

Query Match
Best Local Similarity 1.4%; Score 18; DB 4; Length 3247;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 669 AGAAGTATCTCTGAGAA 686
Db 1443 AGAAGTATCTCTGAGAA 1426

RESULT 24
US-09-620-312D-240/c
; Sequence 240, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PL_FL_genes Version 1.0
; SEQ ID NO 240
; LENGTH: 4953
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4953)
US-09-620-312D-240

Query Match
Best Local Similarity 1.4%; Score 18; DB 4; Length 4953;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 874 CAGCTTAATGTCCTCAAT 891
Db 4194 CAGCTTAATGTCCTCAAT 4177

RESULT 25
US-08-956-171B-143/c
; Sequence 143, Application US/08956171B
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; APPLICANT: Gil H. Choi
; APPLICANT: Patrick S. Dillon
; APPLICANT: Craig A. Rosen
; APPLICANT: Steven C. Barash
; APPLICANT: Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 143:
SEQUENCE CHARACTERISTICS:
LENGTH: 10320 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 143:
US-08-956-171E-143

Query Match 1.4%; Score 18; DB 4; Length 10320;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 GACTTATATTAAAGCTT 965
DB 5759 GACTTATATTAAAGCTT 5742

RESULT 26
US-08-781-986A-143/C
Sequence 143, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunach
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 143:
SEQUENCE CHARACTERISTICS:

LENGTH: 10320 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
US-08-781-986A-143

Query Match 1.4%; Score 18; DB 4; Length 10320;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 948 GACTTATATTAAAGCTT 965
DB 5759 GACTTATATTAAAGCTT 5742

RESULT 27
US-09-790-988-1/C
Sequence 1, Application US/09790988
Patent No. 6632935
GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: MATANABE, HIDEMI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 1.4%; Score 18; DB 4; Length 640681;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 AACTATATTCATCGTTT 278
DB 194919 AACTATATTCATCGTTT 194902

RESULT 28
US-08-016-863-12/C
Sequence 12, Application US/08016863
Patent No. 5602003
GENERAL INFORMATION:
APPLICANT: Pierce, J. M.
APPLICANT: Shorishah, Mohamed G.
APPLICANT: Adler, Beverly
APPLICANT: Fregien, Nevils L.
TITLE OF INVENTION: N-Acetylglucosaminyltransferase V
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/016,863

FILING DATE: 19930210
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/905,795
FILING DATE: 29-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Feider, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 34-92A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)499-8080
TELEFAX: (303)499-8089
TELEX: 823189
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (other)
HYPOTHETICAL: NO
US-08-016-863-12

Query Match 1.3%; Score 17; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 334 CCTGATATCATTCACA 350
DB 24 CCTGATATCATTCACA 8

RESULT 29
US-08-276-968A-12/C
Sequence 12, Application US/08276968A
Patent No. 6015701
GENERAL INFORMATION:
APPLICANT: Pierce, James M.
APPLICANT: Shoreibah, Mohamed G.
APPLICANT: Adler, Beverly L.
APPLICANT: Fregien, Nevis L.
TITLE OF INVENTION: N-Acetylglucosaminyltransferase V
TITLE OF INVENTION: Proteins and Sequences
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,968A
FILING DATE: 19-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/905,795
FILING DATE: 29-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/016,863
FILING DATE: 10-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Feider, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 34-92D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)-499-8080
TELEFAX: (303)-499-8089

TELEX: 49617824
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: YES
US-08-276-968A-12

Query Match 1.3%; Score 17; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 334 CCTGATATCATTCACA 350
DB 24 CCTGATATCATTCACA 8

RESULT 30
US-08-956-171E-1400/C
Sequence 1400, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
APPLICANT: Gil H. Choi
APPLICANT: Patrick S. Dillon
APPLICANT: Craig A. Rosen
APPLICANT: Steven C. Barash
APPLICANT: Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 1400:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1400:
US-08-956-171E-1400
Query Match 1.3%; Score 17; DB 4; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;


```
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2027

Query Match
Best Local Similarity 1.3%; Score 17; DB 3; Length 279;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 ATTTCTTACCATTTCTC 409
Db 161 ATTTCTTACCATTTCTC 145

RESULT 35
US-09-248-796A-10735
; Sequence 10735, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 10735
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-10735

Query Match
Best Local Similarity 1.3%; Score 17; DB 4; Length 291;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 705 AAAATGATTGAAGTA 721
Db 240 AAAATGATTGAAGTA 256

RESULT 36
US-09-540-236-1526/C
; Sequence 1526, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1526
; LENGTH: 348
; TYPE: DNA
; ORGANISM: M.catarrahalis
US-09-540-236-1526

Query Match
Best Local Similarity 1.3%; Score 17; DB 4; Length 348;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 301 GCATTGATCTTGAAT 317
Db 52 GCATTGATCTTGAAT 36

RESULT 37
US-09-270-767-1760/C
; Sequence 1760, Application US/09270767
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1064
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1760
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1760

Query Match
Best Local Similarity 1.3%; Score 17; DB 4; Length 404;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 834 GAAAAGGAAAAGCCAG 850
Db 45 GAAAAGGAAAAGCCAG 29

RESULT 38
US-09-270-767-17042/C
; Sequence 17042, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17042
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17042

Query Match
Best Local Similarity 1.3%; Score 17; DB 4; Length 404;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 834 GAAAAGGAAAAGCCAG 850
Db 45 GAAAAGGAAAAGCCAG 29

RESULT 39
US-09-134-001C-1064
; Sequence 1064, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1064
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1064
```


Query Match 1.3%; Score 17; DB 3; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 956 TTAAGAGTAGATCAT 972
Db 158 TTAAGAGTAGATCAT 174

RESULT 40

US-09-270-767-9605
; Sequence 9605, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-084
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 1999-03-17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9605
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-9605

Query Match 1.3%; Score 17; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 415 ATTACCGAAGTGAAAA 431
Db 257 ATTACCGAAGTGAAAA 273

Search completed: January 26, 2005, 10:01:59
Job time : 117 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 05:47:53 ; Search time 552 Seconds
(without alignments)
13521.574 Million cell updates/sec

Title: US-10-069-062-6

Perfect score: 1299
Sequence: 1 atgtcaaaagcatcttagtgc.....aagactataggtttataa 1299

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1286.2	99.0	1299	US-10-032-585-6023	Sequence 6023, Ap
2	132	10.2	1356	US-09-918-740-48	Sequence 48, Appl
3	132	10.2	1356	US-10-006-909-5	Sequence 5, Appl
4	132	10.2	1356	US-10-369-493-25596	Sequence 25596, A
5	132	10.2	1356	US-10-411-066-5	Sequence 5, Appl
6	132	10.2	1356	US-10-450-941-29	Sequence 29, Appl
7	132	10.2	1356	US-10-835-516-48	Sequence 48, Appl
8	132	10.2	1356	US-10-835-516-48	Sequence 513, App
9	132	10.2	4482	US-10-006-909-9	Sequence 9, Appl
10	132	10.2	4482	US-10-411-066-9	Sequence 9, Appl
11	132	10.2	5051	US-10-006-909-12	Sequence 12, Appl
12	132	10.2	5051	US-10-411-066-12	Sequence 12, Appl

13	132	10.2	5563	US-10-006-909-13	Sequence 13, Appl
14	132	10.2	5563	US-10-411-066-13	Sequence 13, Appl
15	132	10.2	7681	US-09-918-740-61	Sequence 61, Appl
16	132	10.2	7681	US-10-835-516-61	Sequence 61, Appl
17	132	10.2	7693	US-09-918-740-58	Sequence 58, Appl
18	132	10.2	7693	US-10-835-516-58	Sequence 58, Appl
19	132	10.2	7695	US-09-918-740-59	Sequence 59, Appl
20	132	10.2	7695	US-10-835-516-59	Sequence 59, Appl
21	132	10.2	8224	US-09-918-740-62	Sequence 62, Appl
22	132	10.2	8224	US-10-835-516-62	Sequence 62, Appl
23	132	10.2	8235	US-09-918-740-60	Sequence 60, Appl
24	132	10.2	8235	US-10-835-516-60	Sequence 60, Appl
25	132	10.2	8400	US-09-918-740-64	Sequence 64, Appl
26	132	10.2	8400	US-10-835-516-64	Sequence 64, Appl
27	132	10.2	9253	US-10-006-909-7	Sequence 7, Appl
28	132	10.2	9253	US-10-411-066-7	Sequence 7, Appl
29	132	10.2	13917	US-09-918-740-72	Sequence 72, Appl
30	132	10.2	13917	US-10-835-516-72	Sequence 72, Appl
31	132	10.2	14623	US-09-918-740-74	Sequence 74, Appl
32	132	10.2	14623	US-10-835-516-74	Sequence 74, Appl
33	132	10.2	14623	US-10-835-516-74	Sequence 74, Appl
34	132	10.2	14623	US-10-835-516-76	Sequence 76, Appl
35	86.6	6.7	1269	US-10-369-493-25849	Sequence 25849, A
36	67.6	5.2	1306	US-10-320-797-2218	Sequence 2218, Ap
37	66	5.1	1356	US-10-128-714-2406	Sequence 2406, Ap
38	65	5.0	65	US-10-032-585-3023	Sequence 3023, Ap
39	62.6	4.8	1455	US-10-128-714-7406	Sequence 7406, Ap
40	62.6	4.8	1509	US-10-128-714-1406	Sequence 1406, Ap
41	62.6	4.8	1509	US-10-128-714-6406	Sequence 6406, Ap
42	62.6	4.8	3508	US-10-128-714-406	Sequence 406, App
43	62.6	4.8	3509	US-10-128-714-5406	Sequence 5406, Ap
44	46.8	3.6	586	US-10-653-047-1245	Sequence 1245, Ap
45	44.6	3.4	1444	US-10-437-963-96128	Sequence 96128, A

ALIGNMENTS

RESULT 1
US-10-032-585-6023
; Sequence 6023, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6023
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-6023

Query Match 99.0%; Score 1286.2; DB 15; Length 1299;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY	1	ATGTCAAAAGCATTTAGTGCACCTGGAAAGCATTTCTGCTGCTGATTTGTTCTT	60
DB	1	ATGTCAAAAGCATTTAGTGCACCTGGAAAGCATTTCTGCTGCTGATTTGTTCTT	60
QY	61	GAGCCATTATGATGCTTATGACAGCATTTGTATCAGATGATGATGATTTATTA	120
DB	61	GAGCCATTATGATGCTTATGACAGCATTTGTATCAGATGATGATGATTTATTA	120
QY	121	CCAAAGAACGATTTGAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGCAAC	180

Db 121 CAAAAAGAACGTTTGAAGATCTAGATCAAAATTTCTTCAACCCCAATTTTGCAAC 180
Qy 181 GGAGATGGGATATTCACATATCATTAATACAGAGAACCCAGAGATTTCAGCACCC 240
Db 181 GGAGATGGGATATTCACATATCATTAATACAGAGAACCCAGAGATTTCAGCACCC 240
Qy 241 ATAAATCCATTTTGAAGCACTATATTCATCGTTTGTGCTTATATTCACCCGACCGA 300
Db 241 ATAAATCCATTTTGAAGCACTATATTCATCGTTTGTGCTTATATTCACCCGACCGA 300
Qy 301 GCATTTGATCTTGAAATCATCTATTACTCAGACCCCTGATATTCACAGAGATCT 360
Db 301 GCATTTGATCTTGAAATCATCTATTACTCAGACCCCTGATATTCACAGAGATCT 360
Qy 361 GAAACCAACATCCCTCGAATGGAAGAAACCTTTCTTACATCTGTCGCAATTAC 420
Db 361 GAAACCAACATCCCTCGAATGGAAGAAACCTTTCTTACATCTGTCGCAATTAC 420
Qy 421 GAAAGTGAAGAACCCGATTAAGTTCATGCGCAGATGATGTCAGTTGTCACAGT 480
Db 421 GAAAGTGAAGAACCCGATTAAGTTCATGCGCAGATGATGTCAGTTGTCACAGT 480
Qy 481 TTATATCCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGCACAGCT 540
Db 481 TTATATCCCATTTTATCCCAATGTTATCAGTACGATTAAGATATTTTGCACAGCT 540
Qy 541 GCACATATTCACATGTTATGCGCAAAAGATAGATGATGCTGGTGTGATGTCAGCT 600
Db 541 GCACATATTCACATGTTATGCGCAAAAGATAGATGATGCTGGTGTGATGTCAGCT 600
Qy 601 GCATTTATGCTGATGTTATGATAGAGATTCACGACGCTTGTATTAATGACGTTT 660
Db 601 GCATTTATGCTGATGTTATGATAGAGATTCACGACGCTTGTATTAATGACGTTT 660
Qy 661 CAGGTTCTAGAAAGTATCTGAGAAAGTCCCAAGAGTTGAAAAATGATGCAAGT 720
Db 661 CAGGTTCTAGAAAGTATCTGAGAAAGTCCCAAGAGTTGAAAAATGATGCAAGT 720
Qy 721 AACTGGGAATTCAAACATGAAAGATGATACCTACATAGGGAATCAATTAATGAGGT 780
Db 721 AACTGGGAATTCAAACATGAAAGATGATACCTACATAGGGAATCAATTAATGAGGT 780
Qy 781 GACGTCAAGGGTCTCAGAAACACCAATTTGATCAAGATCTCCAAATGAAAAAG 840
Db 781 GACGTCAAGGGTCTCAGAAACACCAATTTGATCAAGATCTCCAAATGAAAAAG 840
Qy 841 GAAAAACCAAGAAAGCTCTGTTGTATGACCAAGCTTAATAGTCCAAATTTACGTT 900
Db 841 GAAAAACCAAGAAAGCTCTGTTGTATGACCAAGCTTAATAGTCCAAATTTACGTT 900
Qy 901 ATGAAGGAATTGAGGGAATGCGGAAAAATATGACCTCAGACCAAGACTTATATTA 960
Db 901 ATGAAGGAATTGAGGGAATGCGGAAAAATATGACCTCAGACCAAGACTTATATTA 960
Qy 961 GAGTTAGATCATCTGTTGAGCTTTGACCTGTTGCGATTAAGAAATCAAAAAAGGTTA 1020
Db 961 GAGTTAGATCATCTGTTGAGCTTTGACCTGTTGCGATTAAGAAATCAAAAAAGGTTA 1020
Qy 1021 CAAGCATTAACAAAAATCAGAGGTTCCATTTGAACCTGATGCTCCAAACCAAGTTGTG 1080
Db 1021 CAAGCATTAACAAAAATCAGAGGTTCCATTTGAACCTGATGCTCCAAACCAAGTTGTG 1080
Qy 1081 GACGTTGTCAAGAGATTCCTGCTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
Db 1081 GACGTTGTCAAGAGATTCCTGCTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1140
Qy 1141 GATGCAATAGCTGATTAATGTTGAAAAATCAAGTGGAAATTTTAAAGCAAAAACTCT 1200
Db 1141 GATGCAATAGCTGATTAATGTTGAAAAATCAAGTGGAAATTTTAAAGCAAAAACTCT 1200
Qy 1201 GAAAAATCAGATTTATTTCTAATAGTTTACTGGGTTGATTTGGAAGCAACAGAGGT 1260
Db 1201 GAAAAATCAGATTTATTTCTAATAGTTTACTGGGTTGATTTGGAAGCAACAGAGGT 1260

Qy 1261 GTACTTGAAGAAAAACAGAAAGCTATATAGTTTATTA 1299
Db 1261 GTACTTGAAGAAAAACAGAAAGCTATATAGTTTATTA 1299
RESULT 2
US-09-918-740-48
; Sequence 48, Application US/09918740
; Publication No. US2003003626A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XCI
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 48
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-918-740-48
Query Match 10.2%; Score 132; DB 10; Length 1356;
Best Local Similarity 50.5%; Pred. No. 4,1e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
Qy 3 GTCAAAAGCATTTAGTGCACCTGAAAAAGCATTTCTGCTGGTGAATTTGGTTCTGA 62
Db 9 GTTGAAGCTTCAATGCTCCCGAGGAAAGCGTTACTAGCTGGATTTAGTTTGA 68
Qy 63 GCCAATTAATGATGCTTATGTCAGCATTTGTCATCAGATGATGATGATTAACACC 122
Db 69 TCAAAATATGAAAGATTTGTGTCGATTAACGCAAGAAAGATGCTGAGCCATCC 128
Qy 123 AAAAGAAC-----CAGTTGAAAGATCTGAATCAAAATTTCTTCCACCCCAATTTGC 176
Db 129 TTACGCTTATTCGAAGGCTCTGATTAAGTTTAAGTGCCTGAGAAAGAAACAAATTTTA 188
Qy 177 AAACGGAGATGGGAATATCATATTCATCAATACAG---AGAAACCAAGAAAGTTCA 233
Db 189 AAGATGGAGTGGCTGATACATATTAAGTCTAAAGTGGCTTCATTCCTGTTGATAGG 248
Qy 234 GTACGCAATTAATCCATTTTGAAGCAACTATATTCATGTTTAACTTAATTCACACC 293
Db 249 CGATCTTAAGAACCTTTCATTTGAAGAAAGTTATGCTAAAGTAAAGTAACTTAACCC 308
Qy 294 GACCAAGCATTTGATCTTGAAA-----TCATCATTTTCTCAGACCTGATATCA 344
Db 309 TAAATGACGACCTACTGCAATTAAGAACTTGTTCCTTAATTTCTTCATGATGATGC 368
Qy 345 TTCACAAAGATTAATGAAACCAAGCATCTCGAATGAGAAAAACATTTCTTTACCA 404
Db 369 CTACATTTCTCAGAGAGATAGGTTACGAAATGTCGCAACAGAAATGATGATTTCA 428
Qy 405 TTCTGCTGCATTAACGAAAGTGAAGAACCGGATTAAGTTAGTTCATGCGCAGATTAAGTGC 464
Db 429 TTGCGCAGAAATTTGAAGAAAGTCCCAAAACAGGGCTGGGCTCTCTGCGAGGTTTAACTC 488
Qy 465 AGTTGTGCAACAAGTTTATTAATCCATTTTATCC-----CAATGTTATCAGTAC 515
Db 489 AGTTTAACTACAGCTTTGCTCTCTTTTGTATGACCTGGAATTAATGATAGACA 548
Qy 516 GAATAAAGATTAATTTGCAACAAGTTCACAGATTTGCAATTTGATGCCCCAAAAAGAT 575
Db 549 ATATTAGAGAGTTATTTATTAATTTAGCAACAAGTGTCTATTTCAAGCTCAGGGTAAAT 608
Qy 576 AGGATCTGGGTTGATGTTGCAACTGCAATTTATAGTCTGATTTGATTAAGAAATTTCA 635

Db 609 TGAAGCGGGTTGATGATGCGCGGACACATATGATCTATCATATATAGAAATTTCC 668
Qy 636 GCCAGCTTGTATTAATGACGTGTTTCAGGTTCTAGAAAGATCCGAGAAGTTCCCA 695
Db 669 ACCCGCATTAATCTCTAATTTGCGAGATATTT-----GGAAGTCTACTTACCGCGTA 722
Qy 696 AGAGTTGAAAAAATTGATGAAAGTAATCTGGGAATTCAAATGAAAGATGATCAATTAC 755
Db 723 ACTGGGGCATTTGGTGTATGAGAGACCTGGAAATATATGATTAAGTAACCATTTAC 782
Qy 756 ATACGGAATCAAGTTATTAATGGGTGACGTCAAGGGTGCAGAAACCCCAATTTGT 815
Db 783 TTCGGGATTAATCTTATGATGAGCGCATATTAAGAAATGTTCAGAAACAGTAACCTGGT 842
Qy 816 ATCAGAGTACTCCATGGAAGAAAGAAAGCAAGAAAGCTGTTGTATGATGACCA 875
Db 843 CCAGAAAGTAAGAAATTTGGTATGATTCGCTATGCGCAAGAAAGCTGTAAGATATATACAG 902
Qy 876 GCTTAATAGTCCCAATTTACAGTTTATGAAGAAAT 911
Db 903 ACTCATATGCAAAATTTCTAGATTTATGATGACT 938

RESULT 3
US-10-069-909-5
; Sequence 5, Application US/1006909
; Publication No. US20030148479A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: KIM, SEON-MON
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: YOSHIKUNI, YASUO
; APPLICANT: NEWMAN, JACK
; APPLICANT: KHEBENIKOV, ARTEM VALENTINOVICH
; TITLE OF INVENTION: BIOSYNTHESIS OF ISOBENTENYL PYROPHOSPHATE
; FILE REFERENCE: 2000-0007
; CURRENT APPLICATION NUMBER: US/10/069,909
; CURRENT FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Phosphomevalonate kinase nucleotide sequence
US-10-069-909-5

Query Match 10.2%; Score 132; DB 15; Length 1356;
Best Local Similarity 50.5%; Pred. No. 4.1e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
Qy 3 GTCGAAGCATTTAGCACTGACCTGGAAGCAATTTCTGCTGGATGATTTGGTTCTTA 62
Db 9 GTTGAGGCTTCAAGGCCCGAGGAAGCGTTACTAGCTGGGATATTTAGTTTAA 68
Qy 63 GCCAATTTATGATGCTTATGTCACAGCATTTGTCATGACGAATSCATGACGATTTAACACC 122
Db 69 TACAAATATGAAGCATTTGTGTGCGATTTATCGGCAAGATGCACTGTGATCCATCC 128
Qy 123 AAAAGAAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGC 176
Db 129 TTACGGTTTATTCGAAGGGCTGATTAAGTTGAAGTCGGTGAAGTAAGTAACAATTTAA 188
Qy 177 AAACGGAATGGGATATCATATCATCAATATACAG--AAGACCCAGAGAATTTCA 233
Db 189 AGATGGGAGTGGCTGTACCATATTAAGTCTTAAAGTGGCTTCAATTTCTGTTTCAATG 248
Qy 234 GTCACGATTAATTCATTTTATAGGCAACTATATTCATGTTTATGCTTATATTCACACC 293

Db 249 CGATCTAAGAACCCCTTTCATTAAGAAAGATATCGCTAACGATATTAAGCTACTTTAAACC 308
Qy 294 GACCGAAGCATTTGATCTTGAA-----TCATCATTTTCTGACCCGATATATCA 344
Db 309 TAAATGACGACATCTGCAATTAAGAACTTGTTCGTTATGATTTATTTCTGATGATGC 368
Qy 345 TTCACAAAGATATCTGAAGAACCAAGACATCTCGAATGAGAAAGAAATTTCTTTACCA 404
Db 369 CTACATCTCTAGAGAGATAGGTTTACGAACATGCTGGCAACAGAAATTTGATTTTCA 428
Qy 405 TTCTGCTGCTATTAACGGAAGTGAAGAACCCGATTAAGTTCAATGCGCAGATTAATGTC 464
Db 429 TTCGACAGAAATTTGAAGAGTTCCCAAAACAGGGCTGGCTCTCCGACGTTTATGCTAC 488
Qy 465 AGTGTTCGCAAGTTTATTTATCCCATTTTATCC-----CAATGTTATCATGAC 515
Db 489 AGTTTAACTTACAGCTTGCCCTCTTTTGTATGACCTGGAAGAAATATATGACAA 548
Qy 516 GATTAAGATATTTTGCACAGCTTGCAAGATTTGACATTTGATGCCCCAAAAGAT 575
Db 549 ATATAGAGAAATTTATCATATTTTGAACAGATTTGCTCATTTGCAAGCTCAGGTAAT 608
Qy 576 AGGATCGGTTTATGATGTTGCAACTGCAATTTATGCTGATTTATATAGAAATTTCA 635
Db 609 TGAAGCGGGTTTGTATGATGCGCGGACACATATGATCATATATATAGAAATTTCC 668
Qy 636 GCCAGCTTGTATTAATGACGTGTTTCAGGTTCTAGAAAGTGAATCCGGAAGATTTCCCA 695
Db 669 ACCCGATTAATCTCTAATTTTGCAGATATTT-----GGAAGTCTACTTACCGCGTA 722
Qy 696 AGAGTTGAAAAAATTGATGAAAGTAATCTGGAATTCAAACATGAAAGATGATCAATTAC 755
Db 723 ACTGGGCGATTTGGTGTATGAGAAAGACCTGGAATTTATGATTAAGTAACCATTTAC 782
Qy 756 ATACGGAATCAAGTTATTAATGGGTGACGTCAAGGGTGCAGAAACCCCAATTTGT 815
Db 783 TTCGGGATTAATCTTATGATGAGCGCATATTAAGAAATGTTCAAGAAACAGTAACCTGGT 842
Qy 816 ATCAGAGTACTCCATGGAAGAAAGAAAGCAAGAAAGCTGTTGTATGATGACCA 875
Db 843 CCAGAAAGTAAGAAATTTGGTATGATTCGCTATGCGCAAGAAAGCTGTAAGATATATACAG 902
Qy 876 GCTTAATAGTCCCAATTTACAGTTTATGAAGAAAT 911
Db 903 ACTCATATGCAAAATTTCTAGATTTATGATGACT 938

RESULT 4
US-10-369-493-25596
; Sequence 25596, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR APPLICATION NUMBER: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25596
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-25596

Query Match 10.2%; Score 132; DB 15; Length 1356;

Best Local Similarity 50.5%; Pred. No. 4,1e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

```
QY 3 GTCAAAAGCATTATGACCTGCGAAAGACATTTCTTGCTGGTGATATTGGTTCTTGA 62
Db 9 GTTGAGAGCCTTCAGTGGCCCGAGGGAAGCGTTACAGCTGGTGATATTGGTTTAA 68
QY 63 GCCAATTATGATGCTTATGTCAGACATGTCATGACGATGACGATGACGATTAACACC 122
Db 69 TACAAATATGAGCATTGTTGATGCGATTTATCGGAAAGATGATGCTGTGATCCATCC 128
QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTGC 176
Db 129 TTAAGGTTTATGCAAGGCTGATGAAAGTTTGAAGTCGTGCAAAAGTAAACAAATTAA 188
QY 177 AAACGAGATGGGAATATCATATCATCAAAATACAG--AGAACCCAGAGAAATTGA 233
Db 189 AGATGGGAGTGGCTGTACATATAGTCTTAAAGTGGCTTCATTCTGTTTCATAG 248
QY 234 GTACAGCATAAATCCATTTTATGAGGCACTATATTATGCTTTTATGCTTATATCAACC 293
Db 249 CGGATTAAGAACCTTTTATGAAAGATTAATGCTAAGTATTTAGCTACTTTAAACC 308
QY 294 GACCGAAGCATTGATCTTGAAA-----TCATCATTTACTCAGACCCCTGATATCA 344
Db 309 TAACATGACGACTACTGCAATGAAATCTTGTTGTTATTTGATTTCTCTGATGATGC 368
QY 345 TTCAACAAGAGTACTGAAACCAAGCATCTCTGAATGAGAGAAAAACATTTCTTTACA 404
Db 369 CTACCAATCTCAGAGAGATAGCGTTTACCAACATCGTGCACACAGAAATGAGTTTCA 428
QY 405 TTCTCGTGCATTACCGAAGTGGAAGAACCCGATAGGTTTCACTCGGCGAGATTAGTGC 464
Db 429 TTTCGACAGAAATTTGAAGAACTTCCCAAAACAGGCTGGGCTCTCGGCGAGTTTATGC 488
QY 465 AGTTGTGCGACAAAGTTTATATCCCATTTTATCCC-----CAATGTTATCAGTAC 515
Db 489 AGTTTAACTACAGCTTTGGCCCTCTTTTGTATGAGCATGGAATAATATGATACAA 548
QY 516 GAATTAAGATTTTTCACACAGCTTGCACAGATTGCACTGTTATGCCCAAAAAAGAT 575
Db 549 ATATAGAGAAGTATTCATATTTAGCACAAAGTGTCTCATTTGCAAGCTCAGGGTAAAT 608
QY 576 AGGATCTGGGTTTATGATGTCACACTGCAATTTATGCTGATGTATATGAGATTTC 635
Db 609 TGGAGCGGTTTGAATGTCAGCGCGGACGATATGATCATGATATGAAAGATTCCC 668
QY 636 GCCAGCTTTGATTAATGACGTGTTTCAGGTTTGAAGATGATCTGAGAAAGTTCCAC 695
Db 669 ACCCGATTAACTCTAATTTGCCAGATAT-----GGAAGTCTACTTACGCGAGTAA 722
QY 696 AAGATTGAAAAAATGATTTGAAGTAACTGCGAATTCMAACATGAAAGATGATATACC 755
Db 723 ACTGGGCGCTTTGGTTGATGAAAGACGTAATATTTACATTTAAAGTAAACCATTTACC 782
QY 756 ATACGAATCAAGTTATTTATGAGGTGACGTCAGAGGTGCTCAGAAACCCCAATTGGT 815
Db 783 TTTCGATTAATCTTATGATGATGCGGCGATTTAAGATGTTCCAGAAACAGTAAACTGG 842
QY 816 ATCAGAGTATCTCAATGCAAAAAAGAAAAAGCAAGAAAGCTCTGTTGTATGACCA 875
Db 843 CCAGAGGTAAAAAAATTTGATATGATTCGATATGCGAAGAAAGCTTGAAATATATACGA 902
QY 876 GCTTAATAGTCCCAATTTACAGTTTATGAAAGAAATT 911
Db 903 ACTGATCATGCAAAATTTAGATTTATGATGATGACT 938
```

RESULT 5
US-10-411-066-5
; Sequence 5, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:

```
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDOR T.
; APPLICANT: NEMMAN, JACK
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-0007.20
; CURRENT APPLICATION NUMBER: US/10/411,066
; CURRENT FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006,909
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-411-066-5
```

Query Match 10.2%; Score 132; DB 16; Length 1356;

Best Local Similarity 50.5%; Pred. No. 4,1e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

```
QY 3 GTCAAAAGCATTATGACCTGCGAAAGACATTTCTTGCTGGTGATATTGGTTCTTGA 62
Db 9 GTTGAGAGCCTTCAGTGGCCCGAGGGAAGCGTTACAGCTGGTGATATTGGTTTAA 68
QY 63 GCCAATTATGATGCTTATGTCAGACATGTCATGACGATGACGATTAACACC 122
Db 69 TACAAATATGAGCATTGTTGATGCGATTTATCGGAAAGATGATGCTGTGATCCATCC 128
QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTGC 176
Db 129 TTAAGGTTTATGCAAGGCTGATGAAAGTTTGAAGTCGTGCAAAAGTAAACAAATTAA 188
QY 177 AAACGAGATGGGAATATCATATCATCAAAATACAG--AGAACCCAGAGAAATTGA 233
Db 189 AGATGGGAGTGGCTGTACATATAGTCTTAAAGTGGCTTCATTCTGTTTCATAG 248
QY 234 GTACAGCATAAATCCATTTTATGAGGCACTATATTATGCTTTTATGCTTATATCAACC 293
Db 249 CGGATTAAGAACCTTTTATGAAAGATTAATGCTAAGTATTTAGCTACTTTAAACC 308
QY 294 GACCGAAGCATTGATCTTGAAA-----TCATCATTTACTCAGACCCCTGATATCA 344
Db 309 TAACATGACGACTACTGCAATGAAATCTGTTGTTATTTGATTTCTCTGATGATGC 368
QY 345 TTCAACAAGATACTGAAACCAAGCATCTCTGAATGAGAGAAAAACATTTCTTTACA 404
Db 369 CTACCAATCTCAGAGAGATAGCGTTTACCAACATCGTGCACACAGAAATGAGTTTCA 428
QY 405 TTCTCGTGCATTACCGAAGTGGAAGAACCCGATAGGTTTCACTCGGCGAGATTAGTGC 464
Db 429 TTTCGACAGAAATTTGAAGAACTTCCCAAAACAGGCTGGGCTCTCGGCGAGTTTATGC 488
QY 465 AGTTGTGCGACAAAGTTTATATCCCATTTTATCCC-----CAATGTTATCAGTAC 515
Db 489 AGTTTAACTACAGCTTTGGCCCTCTTTTGTATGAGCATGGAATAATATGATACAA 548
QY 516 GAATTAAGATTTTTCACACAGCTTGCACAGATTGCACTGTTATGCCCAAAAAAGAT 575
Db 549 ATATAGAGAAGTATTCATATTTAGCACAAAGTGTCTCATTTGCAAGCTCAGGGTAAAT 608
QY 576 AAGATCTGGGTTTATGATGTCACACTGCAATTTATGCTGATGTATATGAGATTTC 635
Db 609 TGGAGCGGTTTGAATGTCAGCGCGGACGATATGATCTATCAGATATGAAAGATTCCC 668
QY 636 GCCAGCTTTGATTAATGACGTGTTTTCAGGTTTGAAGATGATCTGAGAAAGTTCCAC 695
Db 669 ACCCGATTAACTCTAATTTGCCAGATAT-----GGAAGTCTACTTACGCGAGTAA 722
QY 696 AAGATTGAAAAAATGATTTGAAGTAACTGCGAATTCMAACATGAAAGATGATATACC 755
```

Db 723 ACTGGCGATTGGTTGATGAGAAAGACTGGAAATATTACGATTAAAGTAACCATTTACC 782
Qy 756 ATACGAATCAAGTATTATTAATGAGTGAAGGCTGCAAGGCTGCGAAGAACCCAAATTTGGT 815
Db 783 TTCGGGATTACCTTATGATGAGTGGCGATATTAGAAATGGTTCCGAAACAGTAACCTGGT 842
Qy 816 ATCAGAGTACTCCAAATGAGAAAAAGCAAGAAAGCTGTGTGTATGACCA 875
Db 843 CCAGAAAGTAAATAATTTGGTATGATTCGATATGCGAAGAGCTGAAATATATACAGA 902
Qy 876 GCTTAATAGTCCCAATTTACAGTTTATGAAGAAAT 911
Db 903 ACTCGATCATGCAAAATTTAGATTATGATGAGACT 938

RESULT 6

US-10-450-941-29
; Sequence 29, Application US/10450941
; Publication No. US20040063182A1
; GENERAL INFORMATION:
; APPLICANT: TOYOTA JIDOSHA KABUSHIKI KAISHA
; TITLE OF INVENTION: A METHOD OF PRODUCING PRENYLALCOHOL
; FILE REFERENCE: PH-1444PCT
; CURRENT APPLICATION NUMBER: US/10/450,941
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: JP2000-401701
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP2000-403067
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP2001-282978
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: *Saccharomyces cerevisiae*
US-10-450-941-29

Query Match 10.2%; Score 132; DB 16; Length 1356;
Best Local Similarity 50.5%; Pred. No. 4,1e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
Qy 3 GTCAAAGCATTTAGTGAAGCAAGCAATTTCTTGCTGGTGAATTTGGTTCTTGA 62
Db 9 GTTGAGAGCTTCAGTGGCCCGAGGAAAGCGTTACTAGCTGGGATATTTAGTTTAA 68
Qy 63 GCCAATTTATGATGCTTATGTCAGCAGATTGTCATGCAAGATGATGATTAACACC 122
Db 69 TACAAATATGAGCATTTGTAGTCGATTTATCGGCAAGATGATGCTGATGCCATCC 128
Qy 123 AAAAGAAC-----CAGTTGAAAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGC 176
Db 129 TTAAGGTTTATGCAAGGCTGATTAAGTTGAAGCGGTGTAAGAAATTAATTTAA 188
Qy 177 AAACGAGAGATGGGAATATCATATCATCAATAATACG---AGAGCCGAGAGATGA 233
Db 189 AGATGGGAGTGGCTTACATATTAAGTCTTAAGAGTGGCTTATCTGTTTCAATAG 248
Qy 234 GTTACGCAATTAATTCATTTTATAGGCAACTATATTCATGTTTATGCTTATATTCAC 293
Db 249 CGGATCTAAGAACCTTTATGTAAGAAAGTATCGCTAAGTATTTAGCTTAAAC 308
Qy 294 GACCGAAGCATTTGATCTTGAA-----TCATCATTTATCTGAGACCTTGGATATCA 344
Db 309 TAACATGAGCACTACTGCAATTAAGAACTTGTCGTTATTAATTTTCTGTATGATGC 368
Qy 345 TTCACAAGAGATATGAAACCAAGCATCTCGAATGAGAAAGAAATTTCTTTACCA 404
Db 369 CTACCATTTCTCAGAGAGATAGCTTACCAACATCTGTGCAAGAAAGATTTAGTTTCA 428
Qy 405 TTCTGTGCTGCAATTAAGCAAGTGAAGAAAGCCGATTAAGTTTCAATCGGACAGATTAAGTGC 464

Db 429 TTCGACAGAAATTAAGAAAGTTCCAAAACAGGGCTGGGCTCTCGGACAGTTTACTAC 488
Qy 465 AGTTGTGCGCAAGATTATTAATCCCATTTTATCC-----CAATGTTATCAGTAC 515
Db 489 AGTTTAACTACAGCTTTGGCCCTCTTTTGTATCGACCTGAAATTAATATGTAACA 548
Qy 516 GAATTAAGATATTTTGCACACAGTGCACAGATTGCAATTTGATGCCCCAAAAGAT 575
Db 549 ATATAGAGAAATTAATCATTAATTTAGACAAAGTTGCTCATTTGCAAGCTCAGGGTAAAT 608
Qy 576 AGATCTGGGCTTGTATGATGCAATGCAATTTATGCTGATTTATTAAGAAATTTCA 635
Db 609 TGAAGCGGTTTGTATGATGAGCGCGGACAGATATGATCTATCAGATTAAGAAATTTCC 668
Qy 636 GCAGCTTGTATAATAGACGTGTTTCAGGTTCTAGAAATGATTCCTGAGAAATTTCCAC 695
Db 669 ACCCGCATTAATCTCTAATTTGCGAGATTT-----GGAATGCTACTTACGCGATTA 722
Qy 696 AGAGTTGAAAAAATGATTAAGTAAGTAACTGGGAATTCAAACATGAAGATTAATTTACC 755
Db 723 ACTGGCGATTGGTTGATGAGAAAGACTGGAATATTAAGATTAAAGTAACCATTTACC 782
Qy 756 ATACGAATCAAGTATTATTAATGAGTGAAGTGAAGGCTGCAAGGCTGCAAGAAACCCAAATTTGGT 815
Db 783 TTCGGGATTACCTTATGATGATGAGTGGCGATTTAAGATGTTTCAAGAAACAGTAACCTGGT 842
Qy 816 ATCAGAGTACTCCAAATGAGAAAAAGCAAGCAAGAAAGCTGTGTATGACCA 875
Db 843 CCAGAAAGTAAATAATTTGGTATGATTCGATATGCGAATGCGAAGAGCTTAAATATATACAGA 902
Qy 876 GCTTAATAGTCCCAATTTACAGTTTATGAAGAAAT 911
Db 903 ACTCGATCATGCAAAATTTAGATTATGATGAGACT 938

RESULT 7

US-10-835-516-48
; Sequence 48, Application US/10835516
; Publication No. US20040194162A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XCI
; CURRENT APPLICATION NUMBER: US/10/835,516
; PRIOR FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/918,740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 48
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: *Saccharomyces cerevisiae*
US-10-835-516-48

Query Match 10.2%; Score 132; DB 18; Length 1356;
Best Local Similarity 50.5%; Pred. No. 4,1e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
Qy 3 GTCAAAGCATTTAGTGAAGCAAGCAATTTCTTGCTGGTGAATTTGGTTCTTGA 62
Db 9 GTTGAGAGCTTCAGTGGCCCGAGGAAAGCGTTACTAGCTGGGATATTTAGTTTAA 68
Qy 63 GCCAATTTATGATGCTTATGTCAGCAGATTGTCATCAAGATGATGATTAACACC 122
Db 69 TACAAATATGAGCATTTGTAGTCGATTTATCGGCAAGAAATGATGCTGATGCCATCC 128
Qy 123 AAAAGAAC-----CAGTTGAAAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGC 176

Db 129 TTACGGTTTCATGCGAAGGTCGTGATTAAGTTGAAGTCGTGTGAAAAAGTAACAATTAA 188
 QY 177 AAACGAGAAATGGGAATATCATATCATCAAAATACAG---AGAGCCCGAGAGATCA 233
 Db 189 AGATGGGAGTGGCTGTACATATTAAGTCTTAAAGTGGCTTCATCTCTGTTGAGG 248
 QY 234 GTCAAGCAATTAATTCATTTTGAAGCACTATATTCATGCTTTTAACTTATATTCAC 293
 Db 249 CGGATCTAAGAACCTTTCAATTGAAAAAGTATCGTAAGCTTATTAAGTACTTAAAC 308
 QY 294 GACCGAAGCATTTGATCTTAAA-----TCATCTTACTCAGACCTTGATATCA 344
 Db 309 TAAACATGACGACTACTGCAATGAAACTTGTCGTTATGAAATATTTCTCTGATGTC 368
 QY 345 TTCAAGAAATGATCTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTACCA 404
 Db 369 CTACCAATCTCAGAGAGATGCGTTACCGAATCGTGGCAGACAGAAATTTGATTTCA 428
 QY 405 TTCTCGTCCATTAACGAAATGAAAAAGACCGGATTAAGTTTCATCGCAGATTAAGTGC 464
 Db 429 TTGCGACAGAAATGAAGAAAGTCCCAAAACAGGGCTGGGCTCCTCGGAGGTTTATGTCAG 488
 QY 465 AGTTGTTGCCAAGTTTATATCCATTTTATCC-----CAATGTTATCAGTAC 515
 Db 489 AGTTTAACTACAGCTTTGGCTCTTTTGTATGGAACCTGAAAAATATGTAAGCA 548
 QY 516 GAATTAAGATATTTTGCACACGTTGCACAGATTGTCATGTTATGCCAAAAAAGAT 575
 Db 549 ATATAGAAAGTATTAATCAATATTAAGCAAGTTGCTCATTTGCAAGCTCAGGGTAAAT 608
 QY 576 AGGATCTGGGTTGATGTTGCAACTGCAATTATATGATGTAATATGAAAGATTCA 635
 Db 609 TGAAGACGGGTTGATGATGACGGCGGAGCATATGATCATATATAGAAAGATTCC 668
 QY 636 GCCAGCTTGAATTAAGAGTGTTCAGGTTCTAGAAAGTGAATCCGAAAGTTCGCCAC 695
 Db 669 ACCCGATTAATCTTAATTTGCCAATATT-----GAAAGTCTACTTAACCGCATTA 722
 QY 696 AGAGTTGAAAAAATGATTGAAGTAAGTAAGTGAATTAAGTAAGTAAGTATTAAC 755
 Db 723 ACTGGGCAATTTGGTATGAAAGAACTGGAATATTAAGATTAAAGTAAACATTTTAC 782
 QY 756 ATACGAATCAAGTTTATATGAGGTGACGTCAGAGGTGCTGAGAAACCCAAATGCT 815
 Db 783 TTGGGATTAATCTTAATGATGAGCGATATTAAGATGTTCAAGAAACGTAACCTGCT 842
 QY 816 ATCACGACTCTCAATGAAAAAAGAAAGCCAGAAAGGCTGTTGTATGACCA 875
 Db 843 CCAAGAGTAAATAATTTGGTATGATTCGATATGCCAGAAAGCTTGAATAATATACGA 902
 QY 876 GCTTAATAGTCCATTTTACAGTTTATGAAGAAAT 911
 Db 903 ACTGATCATGCAAAATTTAGATTTATGATGAGTACT 938
 RESULT 8
 US-10-793-639-513
 ; Sequence 513, Application US/10793639
 ; Publication No. US20040199940A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Karunananda, Balasubjini
 ; APPLICANT: Yu, Jaehyuk
 ; APPLICANT: Kishore, Ganesh M.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
 ; FILE REFERENCE: 16516.075
 ; CURRENT APPLICATION NUMBER: US/10/793,639
 ; PRIOR APPLICATION NUMBER: US/09/614,221A
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: US 60/142,981
 ; PRIOR FILING DATE: 1999-07-12
 ; NUMBER OF SEQ ID NOS: 626

; SEQ ID NO 513
 ; LENGTH: 1356
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces cerevisiae
 US-10-793-639-513
 Query Match 10.2%; Score 132; DB 18; Length 1356;
 Best Local Similarity 50.5%; Pred No. 4.1e-23;
 Matches 4/3; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
 QY 3 GTCAAAAGCATTTAGTGAACCTGAAAAAGCATTTTCTGTGTGATATTTGGTCTTGA 62
 Db 9 GTTGAAGGCTTCAAGTCCCGAGGAAAGCGTTACTTACCTGTGATATTTAGTTTGA 68
 QY 63 GCCATTTATGATGTTATGACAGCAATTTGCATCAGAAATGATGATGCAATTAACAC 122
 Db 69 TACAAATATGAAAGATTTGATGATGATTTTCGCAAGAAATGCAATGCTGAGCCATCC 128
 QY 123 AAAAGAAC-----CAGTTGAAGAAATCTAGAAATCAAAATTTCTCACCCCAATTTGC 176
 Db 129 TTACGGTTTCATGCAAGGCTCTGATTAAGTTGAAAGTGGCTGAAAAAGTAACAATTAA 188
 QY 177 AAACGAGAAATGGGAATATCATATCATCAAAATACAG---AGAGCCCGAGAGATCA 233
 Db 189 AGATGGGAGTGGCTGTACATATTAAGTCTTAAAGTGGCTTCATCTCTGTTGATGAG 248
 QY 234 GTCAAGCAATTAATTCATTTTGAAGCACTATATTCATGCTTTTAACTTATATTCAC 293
 Db 249 CGGATCTAAGAACCTTTCAATTGAAAAAGTATGCTTAACGTAATTTAGCTACTTAAAC 308
 QY 294 GACCGAAGCATTTGATCTTAAA-----TCATCTTACTCAGACCTTGATATCA 344
 Db 309 TAAACATGACGACTACTGCAATGAAACTTGTCGTTATGAAATATTTCTCTGATGTC 368
 QY 345 TTCAAGAAATGATCTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTACCA 404
 Db 369 CTACCAATCTCAGAGAGATGCGTTACCGAATCGTGGCAACAGAAATGTAAGTTC 428
 QY 405 TTCTCGTCCATTAACGAAATGAAAAAGACCGGATTAAGTTTCAATCGCAGATTAAGTGC 464
 Db 429 TTGCGACAGAAATGAAGAAAGTCCCAAAACAGGGCTGGGCTCCTCGGAGGTTTATGTCAG 488
 QY 465 AGTTGTTGCCAAGTTTATATCCATTTTATCC-----CAATGTTATCAGTAC 515
 Db 489 AGTTTAACTACAGCTTTGGCTCTTTTGTATGGAACCTGAAAAATATGTAAGCA 548
 QY 516 GAATTAAGATATTTTGCACACGTTGCACAGATTGTCATGTTATGCCAAAAAAGAT 575
 Db 549 ATATAGAAAGTATTAATCAATATTAAGCAAGTGTCTCATTTGTCAAGCTCAGGGTAAAT 608
 QY 576 AGGATCTGGGTTGATGTTGCAACTGCAATTATATGATGTAATATGAAAGATTCA 635
 Db 609 TGAAGACGGGTTGATGATGACGGCGGAGCATATGATCTATGCAATATGAAAGATTCC 668
 QY 636 GCCAGCTTGAATTAAGAGTGTTCAGGTTCTAGAAAGTGAATCCGAAAGTTCGCCAC 695
 Db 669 ACCCGATTAATCTTAATTTGCCAATATT-----GAAAGTCTACTTAACCGCATTA 722
 QY 696 AGAGTTGAAAAAATGATTGAAGTAAGTAAGTGAATTAAGTAAAGATGTAACCT 755
 Db 723 ACTGGGATTAATCTTAATGATGAGCGATATTAAGATGTTCAAGAAACGTAACCTGCT 782
 QY 756 ATACGAATCAAGTTTATATGAGGTGACGTCAGAGGTGCTGAGAAACCCAAATGCT 815
 Db 783 TTGGGATTAATCTTAATGATGAGCGCATTAATGAAATGTTCAAGAAACGTAACCTGCT 842
 QY 816 ATCACGACTCTCAATGAAAAAAGAAAGCCAGAAAGGCTGTTGTATGACCA 875
 Db 843 CCAAGAGTAAATAATTTGGTATGATTCGATATGCCAGAAAGCTTGAATAATATACGA 902
 QY 876 GCTTAATAGTCCATTTTACAGTTTATGAAGAAAT 911
 Db 903 ACTGATCATGCAAAATTTAGATTTATGATGAGTACT 938

RESULT 9
US-10-006-909-9
Sequence 9, Application US/10006909
Publication No. US20030148479A1
GENERAL INFORMATION:
APPLICANT: KEASLING, JAY
APPLICANT: MARTIN, VINCENT
APPLICANT: PITERA, DOUGLAS
APPLICANT: KIM, SEON-WON
APPLICANT: WITHERS III, SYDNOR T.
APPLICANT: YOSHIKUNI, YASUO
APPLICANT: NEWMAN, JACK
APPLICANT: KHEBENIKOV, ARTEM VALENTINOVICH
TITLE OF INVENTION: BIOSYNTHESIS OF ISOENTENYL PYROPHOSPHATE
FILE REFERENCE: 2000-0007
CURRENT APPLICATION NUMBER: US/10/006, 909
CURRENT FILING DATE: 2002-04-02
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 4482
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: "MEVB" operon nucleotide sequence
US-10-006-909-9

Query Match 10.2%; Score 132; DB 15; Length 4482;
Best Local Similarity 50.5%; Pred. No. 7.8e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTGCTGGTGGATTTGGTTCTTGA 62
DB 1606 GTTGAAGGCTTCAGAGCCCGCAGGAAAGCGTACAGCTGGGATATTAGTTTAA 1665
QY 63 GCCAATTATGATGCTTATGTGACAGCATTTGCATGACGAATGACGAGTTAAACACC 122
DB 1666 TACAAATATGAGCATTTGTGATCGAATTATGCGAAGATGATGCTGTACCCATCC 1725
QY 123 AAAAGAAC-----CAGTTTGAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTTGC 176
DB 1726 TTAAGGTTTATGCAAGGGCTGATTAAGTTGAAGGCGGTGAAAGTAAACATTTAA 1785
QY 177 AAACGAGATGGGAATATCATCATCAATTAACAG--AGAAGCCAGAGAATTGA 233
DB 1786 AGATGGGAGTGCTGTACCATTAAGTCTTAAAGTGGCTTCATTCGTTTGAATAG 1845
QY 234 GTACGCGATTAATCCATTTTATGAGGCACTATATTCATCGTTTATGCTTATTAAC 293
DB 1846 CGGATCTAAGAACCTTTATGAAAAGTTATCGCTAACGATTTAGCTACTTTAAAC 1905
QY 294 GACCGAAGCATTTGATCTTGA--TCACTATTACTAGACCTCGATATCA 344
DB 1906 TAAACATGACGACTACTGCAATGAAACTTGTGTTATGATTAATTTCTGTATGATG 1965
QY 345 TTCAACAAGAAATGCTGAAACCAAGACATCTCGAATGAGAAAACATTTCTTACCA 404
DB 1966 CTACCACTTCTGAGAGATAGCTTACCGAACATCTGTGCAACAGAAATGTGTTTCA 2025
QY 405 TTCTGTGTCATTAAGGAGTGAAGAACCCGATTAAGTTCACTGCGAGATTAAGTGC 464
DB 2026 TTTCGACAGATTAAGAAAGTTCCTCAAAACAGGCTGGGCTCTCGCAGGTTTATGAC 2085
QY 465 AGTTGTGCGACAAAGTTTATATCCATTTTATCCC-----CAATGTTATCAATAC 515
DB 2086 AGTTTAACTACAGCTTTGGCTCTTTTGTATGAGACCTGGAATAATATGATAGCA 2145
QY 516 GAATTAAGATTTTGTGACAACTGTGACAGATTTGACATGTTATGCCCCAAAAAGAT 575
DB 2146 ATATAGAGAAATTAATCATTAATTTAGACAAAGTTGCTCATTTCAAGCTCAGGTTAAAT 2205

QY 576 AGGATCTGGGTTTGTATGTTGCAACTGCAATTAATGCTGTGATTAATAGAAATTTCA 635
DB 2206 TCGAAGCGGTTTGTATGTCAGCGCCGACGATATGATCTATCAGATATAGAAATTTCC 2265
QY 636 GCCAGCTTGTATTAATGACGTGTTCAAGTTCTAGAAATGATCTCGAAGATTTCCCA 695
DB 2266 ACCCGATTAATCTCTAATTTTCCAGATATTT-----GGAAAGCTACTTACGCGATTA 2319
QY 696 AGAGTTGAAAAATTTGATGAAGTAAGTAATCTGGAATTCACATGAAAGATGTACATTACC 755
DB 2320 ACTGGGCGCTTTGGTTGATGAAAGACCTGGAATATTAAGATTAAGATACATTACC 2379
QY 756 ATACGGAATCAATTATTAATGAGTGCATCAAGGGTGCAGAAACCCCAATTTGAT 815
DB 2380 TTGGGATTAATCTTATGATGAGGCGATATTAAGAAATGTTGAGAAACAGTAAACTGG 2439
QY 816 ATCAGAGTACTTCCAAATGGAAGAAAGCAAGAAAGCTCTGTTGTATGACCA 875
DB 2440 CCAGAAAGTAAATAATTTGATGATTCGCATATGCCAGAAAGCTTGAATAATATACAG 2499
QY 876 GCTTAATAGTGCATTTACAGTTTATGAAGGATTT 911
DB 2500 ACTGATCATGCAAAATTTAGATTAATGATGACT 2535

RESULT 10
US-10-411-066-9
Sequence 9, Application US/10411066
Publication No. US20040005678A1
GENERAL INFORMATION:
APPLICANT: KEASLING, JAY
APPLICANT: MARTIN, VINCENT
APPLICANT: PITERA, DOUGLAS
APPLICANT: WITHERS III, SYDNOR T.
APPLICANT: NEWMAN, JACK
TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
FILE REFERENCE: 2000-0007.20
CURRENT APPLICATION NUMBER: US/10/411, 066
CURRENT FILING DATE: 2000-04-09
PRIOR APPLICATION NUMBER: 10/006, 909
PRIOR FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 124
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 4482
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: "MEVB" operon nucleotide sequence
US-10-411-066-9

Query Match 10.2%; Score 132; DB 16; Length 4482;
Best Local Similarity 50.5%; Pred. No. 7.8e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTGCTGGTGGATTTGGTTCTTGA 62
DB 1606 GTTGAAGGCTTCAGAGCCCGCAGGAAAGCGTACAGCTGGGATATTAGTTTAA 1665
QY 63 GCCAATTATGATGCTTATGTGACAGCATTTGCATGACGAATGACGAGTTTAAACACC 122
DB 1666 TACAAATATGAGCATTTGTGATCGAATTATGCGAAGATGATGCTGTACCCATCC 1725
QY 123 AAAAGAAC-----CAGTTTGAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTTGC 176
DB 1726 TTAAGGTTTATGCAAGGGCTGATTAAGTTGAAGGCGGTGAAAGTAAACATTTAA 1785
QY 177 AAACGAGATGGGAATATCATCATCAATTAACAG--AGAAGCCAGAGAATTGA 233
DB 1786 AGATGGGAGTGCTGTACCATTAATAGTCTTAAAGTGGCTTCATTCGTTTGAATAG 1845

RESULT 12
US-10-411-066-12
; Sequence 12, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: NEWMAN, JACK
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-0007.20
; CURRENT APPLICATION NUMBER: US/10/411, 066
; PRIOR FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006, 909
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 5051
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: "MBI" operon nucleotide sequence
US-10-411-066-12

Query Match 10.2%; Score 132; DB 16; Length 5051;
Best Local Similarity 50.5%; Pred. No. 8,4e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
QY 3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTGTGTGATATTGGTCTTGA 62
DB 1606 GTGAAAGCCTTCAAGTCCCGAGAAAGGCTTACTAGCTGTGATATTAGTTTGA 1665
QY 63 GCCAATTTATGATGCTTATGTGACAGCATTTGCATACGAAATGCAATGCAATTAACCC 122
DB 1666 TACAAATATGAAGCATTTGTAGTCGATATATCGGCAAGATGATGCTTACCCATCC 1725
QY 123 AAAAGAAC-----CAGTTGAAAGATTTAGATCAAAATTTCTTCAACCCCAATTTGC 176
DB 1726 TTACGGTTTATGCAAGGGCTGATTAAGTTGAAGTGGTGTGAAAGATTAACAAATTTAA 1785
QY 177 AAACGAGATGGGAATATCATATCATCAATATACAG--AGAAAGCCAGAGAATTC 233
DB 1786 AGATGGAGATGGCTGTACATATTAAGTCTTAAAGTGGCTTCAATTCCTTTGATAGG 1845
QY 234 GTACGCAATTAATTCATTTTATGAGGCACATATATTCATGTTTATGCTTATATCAACC 293
DB 1846 CGGATCTAAGAACCTTTTCAATGAAAAGTTATCGTAACTGTTAGTAACTTAAACC 1905
QY 294 GACCGAAGCATTTGATCTTGAA-----TCATCTTACTCAGACCTTGATATCA 344
DB 1906 TTAACATGACGACTACTGCAATGAAACTTGTTCGTTATGATATTTCTCTGATGATGC 1965
QY 345 TTCACAAGAACTACTGAACCAAGACATCTCGAATGAGAAAACATTTCTTTACCA 404
DB 1966 CTACCATTTCTCAGAGATAGCTTACCAACATCTGCGAACAAGAAATTTGATTTCA 2025
QY 405 TTCTGTCGCAATTCAGAAATGAAAAGACCGGATTAAGTTTCAATCGGACAGATTAAGTGC 464
DB 2026 TTGCGACAGAAATGAAGAGTTCCCAAAACAGGGCTGGGCTCTCGGCAAGTTTATGAC 2085
QY 465 AGTTGTGCGCAAGATTTATATCCATTTTATCC-----CAATGTTATCAAGTAC 515
DB 2086 AGTTTAACTACAGCTTGGCTCTCTTTTGTATGAGACCTGAAATATATGTAAGCA 2145
QY 516 GAATTAAGATATTTTGCACAAGCTTGACAGATTTGATGATTTTATGCCCCAAAAAAGAT 575
DB 2146 ATATAGAGAAATTTATATTTAGCAAGATTTGCTCATTTGTCAGCTTCAAGGTAAT 2205
QY 576 AGGATCTGGTTGATGTTGCAACTGCAATTTATGCTGATTTATATGAAGATTTCA 635

DB 2206 TGGAAAGCGGTTTATGATGACGGCGGACAGCATATGATCTATCATATAGAAATTTCC 2265
QY 636 GCCAGCTTGAATTAATAGAGTGTTCAGGTTCTAGAAAGTATCTGAGAAATTTCCAC 695
DB 2266 ACCCGCATTAATCTCTAATTTTGGCAGATATTT-----GGAAGTCTACTTACGGCAGTA 2319
QY 696 AGAGTTGAAAAAATGATGAAAGTAACTGGGAATTCAAACATGAAAGATGATTAAC 755
DB 2320 ACTGGCGCATTTGTTGATGAAAGACATGGAATTTATGATTAAGTAACCATTTAC 2379
QY 756 ATACGAATCAAGTTATTAATGAGTACAGTCAAGGTTGCTCGAAGAACCCAAATTTGT 815
DB 2380 TTGCGGATTAATCTTATGATGAGGCGATATTAAGAAATGTTTCGAAACAGTAAACTG 2439
QY 816 ATCAGAGTACTGCATATGAAAAAGAAAAAGCCAGAAAGAGCTGTGTATGACCA 875
DB 2440 CCAGAAAGTAAAAAATTTGATATGATTCGATATGCGATATGCCAAGAGCTTGAATATATACAG 2499
QY 876 GCTTAATAGTCCCATTTTACAGTTTATGAAGAAAT 911
DB 2500 ACTGCATCATGCAAAATTTCAATTTATGATGACT 2535

RESULT 13
US-10-006-909-13
; Sequence 13, Application US/10006909
; Publication No. US20030148479A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: KIM, SEON-WON
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: YOSHIKUNI, YASUO
; APPLICANT: NEWMAN, JACK
; APPLICANT: KHEBNIKOV, ARTEM VALENTINOVICH
; TITLE OF INVENTION: BIOSYNTHESIS OF ISOPENENTYL PYROPHOSPHATE
; FILE REFERENCE: 2000-0007
; CURRENT APPLICATION NUMBER: US/10/006, 909
; CURRENT FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 5963
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: "MBIS" operon nucleotide sequence
US-10-006-909-13

Query Match 10.2%; Score 132; DB 15; Length 5963;
Best Local Similarity 50.5%; Pred. No. 9.1e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
QY 3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTGTGTGATATTGGTCTTGA 62
DB 1606 GTTGAAGCCTTCAAGTCCCGAGAAAGGCTTACTAGCTGTGATATTAGTTTGA 1665
QY 63 GCCAATTTATGATGCTTATGTGACAGCATTTGCATGCAAGATGCAATGATTAACCC 122
DB 1666 TACAAATATGAAGCATTTGTAGTGAATTTATGCAAGATGCAATGCTTACCCATCC 1725
QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGC 176
DB 1726 TTACGGTTTATGCAAGGGCTGATTAAGTTGAAGTGGTGTGAAAGATTAACAAATTTAA 1785
QY 177 AAACGAGATGGGAATATCATATCATCAATATACAG--AGAAAGCCAGAGAATTC 233
DB 1786 AGATGGAGATGGCTGTACATATTAAGTCTTAAAGTGGCTTCAATTCCTTTGATAGG 1845
QY 234 GTACGCAATTAATTCATTTTATGAGCAACTATATTTATGCTTATATTAATTAACCC 293

```

Db      1846 CGGATCTTAAGAACCTTTCATTGAAAAGATTATGCTAAGTATTAGCTACTTTAAACC 1905
Qy      234  GACCGAAGCATTTGATCTTGA---TCATCATTTTACTGAGCCCTGGATATCA 344
Db      1906 TAACTAGGACGACTCTGCAATAGAACTGTCTTATTTGATTTTCTCTGATGATGC 1965
Qy      345  TTCAACAAGATCTGAAACCAAGACATCTCGAATGAGAAAAACATTTCTTTTACA 404
Db      1966 CTACATCTCTGAGGAGTAGCGTTACCGAACATCTGGCAACAGAAATTTAGATTTC 2025
Qy      405  TTCTGTCATTTTCCGAAAGTGAAGAACCGGANTAGTTTCACTCGGAGATTAAGTCTC 464
Db      2026 TTGCGACAGAAATGGAAGAGTTCCCAAAACAGGGCTGGGCTCTCGGAGGTTTACAC 2085
Qy      465  AGTTGTCACCAAGTTTATTATCCATTTTATCCC-----CAATGTTATCAGTAC 515
Db      2086 AGTTTAACTACAGCTTTGGCTCTTTTGTATCGAACCTGGAAATATATGACAA 2145
Qy      516  GAATTAAGATTTTGGCAACGTTGCAACGATTTGACATTTGTTATGCCAAAAAAGAT 575
Db      2146 ATATAGAGAGTTATTTCAATTAATTTAGCAAGTTGCTCATTTGTCACGCTGAGGTA 2205
Qy      576  AGGATCTGGGTTTATGTTGCACTGCAATTTATGTCGATTTGATATATGAAATTTCA 635
Db      2206 TGAAGCGGGTTTATGATGAGCGGCGGACATATGATCTATCATATATGAAATTTCC 2265
Qy      636  GCCAGCTTTGATTAATGACGTGTTCAAGTTCTAGAAAGTATCTGAGAAAGTTCCAC 695
Db      2266 ACCCGCATTAATCTTAATTTGCGAGATTT-----GGAAGTCTACTTACGCGATA 2319
Qy      696  AGAGTTAAAAAATTTGATTTGAAAGTACTGGGAATTCAAACATGAAAGATGATAC 755
Db      2320 ACTGGCCATTTGGTTATGAAAGAACTGAAATTTATGATTAAGATTAACCATTTTAC 2379
Qy      756  ATACGGAATCAAGTTATTAATGCGTGAACGTCAAGGTGCTCAGAAAACCCAAATTTGT 815
Db      2380 TTGGGGATTAATCTTTATGATGCGCATATTAGAAATGTTCAAGAAACGTAATACTGT 2439
Qy      816  ATCAGAGTACTCAATGGAAGAAAGAAAGCAAGAAAGCTCTGTTGTGTATGACCA 875
Db      2440 CCAGAAAGTAAAAAATTTGATGATTCGCAATATGCCAGAAAGCTGAAAAATATAC 2499
Qy      876  GCTTAATAGTCCAAATTTACGTTTATGAAAGAAAT 911
Db      2500 ACTGATCATGCAAAATTTCTGATTTATGATGACT 2535

RESULT 14
US-10-411-066-13
; Sequence 13, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDNOR T.
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-0007.20
; CURRENT APPLICATION NUMBER: US/10/411,066
; PRIOR FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006,909
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 5963
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-10-411-066-13

```

```

Query Match      10.2%; Score 132; DB 16; Length 5963;
Best Local Similarity 50.5%; Pred. No. 9,1e-23;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

Qy      3  GTCAAAAGCATTTTATGACCTGCGAAAGCATTTTCTTGTGTGATATTTGGTCTTGA 62
Db      1606 GTTGAAGCTTCAAGCCCCGAGGAAAGCGTTACTACTGTGTGATATTTAGATTGA 1665
Qy      63  GCCAATTTATGATGCTTATGACAGCATTTGATCAGAAATGACAGTATTAACACC 122
Db      1666 TACAAAATATGAAAGCATTTGATGATGATTAACGCAAGATGATGCTGATGCCATTC 1725
Qy      123  AAAAGAAC-----CAGTTCAAGAAATCTGAATCAAAATTTCTTCAACCCCAATTTGC 176
Db      1726 TTACGTTTATTTGCAAGGCTGTGATTAAGTTGAAAGTGTGAAAAGTAAACATTTTA 1785
Qy      177  AAACGAGAAATGGGAATATCACATTCATCAATACG---AGAAAGCCAGAGAAAGTTCA 233
Db      1786 AGATGGGAGTGTGCTGTACATATTAAGTCTTAAAGTGGCTTCAATCTGTTGATAGG 1845
Qy      234  GTACGCAATTAATCCATTTTATGAGGCACTATATTCATCGTTTATGCTTATATTAAC 293
Db      1846 CGGATCTAAGAACCTTTCAATTGAAAAGTATGCTAAGATTTAGCTACTTTAAACC 1905
Qy      294  GACCGAACATTTGATCTTGA---TCATCATTTACTCAGACCCCTGATATCA 344
Db      1906 TAACTAGGACGACTCTGCAATTAAGAACTGTGTTATGATTAATTTTCTCTGATGATGC 1965
Qy      345  TTCAACAAGATCTGAAACCAAGACATCTCGAATGAGAAAAACATTTCTTTTACA 404
Db      1966 CTACATCTCTGAGGAGTAGCGTTACGAAACATCGTGGCAACAGAAATTTAGATTTC 2025
Qy      405  TTCTGTCATTTTCCGAAAGTGAAGAACCGGANTAGTTTCACTCGGAGATTAAGTCTC 464
Db      2026 TTGCGACAGAAATGGAAGAGTTCCCAAAACAGGGCTGGGCTCTCGGAGGTTTACAC 2085
Qy      465  AGTTGTCACCAAGTTTATTTATCCATTTTATCCC-----CAATGTTATCAGTAC 515
Db      2086 AGTTTAACTACACTTTTGGCTCTTTTGTATGCACTGGAATATATGATGACAA 2145
Qy      516  GAATTAAGATTTTGGCAACGTTGCAAGATTTGACATTTGTTATGCCAAAAAAGAT 575
Db      2146 ATATAGAGAGTATTAATTAATTTATGACCAAGTTGCTCATTTGCAAGCTCAGGTA 2205
Qy      576  AGATCTGGGTTTGAATGTCACATGCAATTTATGCTGATTTATATGAAAGATTTCA 635
Db      2206 TGAAGCGGGTTTATGATGAGCGGCGGACATATGATCTATCAGATTAAGAAATTTCC 2265
Qy      636  GCCAGCTTTGATTAATGACGTGTTCAAGTTCTAGAAAGTATCTGAGAAAGTTCCAC 695
Db      2266 ACCCGCATTAATCTTAATTTGCGAGATTT-----GGAAGTCTACTTACGCGATA 2319
Qy      696  AGAGTTAAAAAATTTGATTTGAAAGTACTGGGAATTCAAACATGAAAGATGATAC 755
Db      2320 ACTGGCCATTTGTTGATGAGAAAGACGGAATATTAAGATTAAGTAAACCATTTTAC 2379
Qy      756  ATACGGAATCAAGTTATTAATGCGTGAACGTGCTGAGAAAGCCCAAAATTTGT 815
Db      2380 TTGGGGATTAATCTTTATGATGCGCATATTAGAAATGTTCAAGAAACGTAATACTGT 2439
Qy      816  ATACAGAGTACTCAATGGAAGAAAGAAAGCCAGAGAAAGCTCTGTTGTATGACCA 875
Db      2440 CCAGAAAGTAAAAAATTTGATGATTCGATATGCAAGAAAGCTGAAAAATATATAC 2499
Qy      876  GCTTAATAGCCAAATTTACGTTTATGAAAGAAAT 911
Db      2500 ACTGATCATGCAAAATTTCTGATTTATGATGACT 2535

RESULT 15
US-09-918-740-61
; Sequence 61, Application US/09918740

```

```

; Publication No. US20030033626A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehnle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; TITLE OF INVENTION: create novel traits in transgenic organisms
; FILE REFERENCE: KAS-103XC1
; CURRENT APPLICATION NUMBER: US/09/918,740
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patencin version 3.0
; SEQ ID NO 61
; LENGTH: 7681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Operon C containing A. thaliana, S. cerevisiae, and Streptomyces
; OTHER INFORMATION: CL190
; OTHER INFORMATION: DNA, and R. capsulatus DNA
US-09-918-740-61

```

```

Query Match      10.2%; Score 132; DB 10; Length 7681;
Best Local Similarity 50.5%; Pred. No. 1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

```

```

QY 3 GTCMAAAGATTGACGACCTGGAAGAAAGCAATTTCTGCTGGTGAATTTGGTTCTTGA 62
DB 43 GTTGAGAGCTTCAGGCCCCAGGGAACGGTTACGTGGTGGATTTTACTTTTAA 102
QY 63 GCCAATTTATGATGCTTATGTGACAGCATTTGTCATGCAAGATGATGATTAACACC 122
DB 103 TACMAAATATGAGCAATTTGTAGTCGATTTATGCGAATATGCAATGCTGTAGCCATCC 162
QY 123 AAAAGAAC-----CAGTTTGAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTGC 176
DB 163 TTACGGTTCATTCAGCAAGGCTGTGATTAAGTTGAAGCGGTGAAAGTAACAAATTTAA 222
QY 177 AAACGAGAAATGGGAATATCATCATCAATCAATACAG--AGAACCCAGAGAAATGCA 233
DB 223 AATATGGGAGTGGCTGTACCATATTAAGTCTTAAAGTGGCTTCATTCCTGTTTCATGAG 282
QY 234 GTCAGCATTAATTCATTTTGAAGGCACTATATTATCATGTTTGAAGTTATTCAC 293
DB 283 CGGATTAAGAACCTTTTCAATGAAAAGTTATCGCTAACGATTTAGCTACTTTAAAC 342
QY 294 GACCGAAGCATTTGATCTTGAA-----TCATCATTTTACTCAGACCCTGGAATATCA 344
DB 343 TAAACATGACGACTACTGCAATAGAAACTTGTTCGTTATTTGATTTTCTGTATGATGC 402
QY 345 TTCACAAAGATATCTGAAAACCAAGCATCTCGAATGAGAAAACAAATTTCTTTACA 404
DB 403 CTACCATTTCTCAGAGATAGAGCTTACCGAACATCTGTGCAACAGAAATGAGTTTCA 462
QY 405 TTCTCGTCATTAACGAGATGGAAGAACCCGATTAAGTTTCTCGGACGAGATTAAGTC 464
DB 463 TTCGCAAGAAATGAAAGATTCCTCAAAACAGGGCTGGGCTCTCGGACGATTTAGTAC 522
QY 465 AGTTGTGCGACAGATTTTATCCCATTTTATCCC-----CAATGTTATCAGTAC 515
DB 523 AGTTTAACTACAGCTTTGGCCCTCTTTTGTATGAGACCTGGAATAATATAGACAA 582
QY 516 GAATTAAGATATTTTGCACACGTTGCACAGATTTGCACATTTGATGCCCCAAAAAAGAT 575
DB 583 ATATAGAGAAGATATTCATTAATTTAGCAACAAGTTGCTCATTTCAAGCTCAGGATAAAT 642
QY 576 AGGATATGAGTTGATGTTGCACTGCAATTTATGCTGATGTTATATAGAGATTTCA 635
DB 643 TGGAGCGGTTTGAATGAGCGCGGACAGATATGATCTATCAAGATATAGAAATTTCC 702
QY 636 GCCAGCTTTGATTAATAGAGTGTTCAGGTTCTAGAAAGTATCTGAGAAATTTCCAC 695

```

```

DB 703 ACCCGCATTAATCTCTAATTTCCGAGATTT-----GGAAGTCTACTTACGGCAGTAA 756
QY 696 AGAGTTGAAAAAATTTGATTAAGAAAGTAACTGGGAATTCAAACATGAAGATGATACCTACC 755
DB 757 ACTGGGCGATTTGGTTGATGAAAGAGCTGGAATTTAGATTTAAAGATACCATTTAC 816
QY 756 ATACGAAATCAAGTTATTAATGAGTGAAGTCAAGGCTGCTAGAAACCCCAATTTGGT 815
DB 817 TTCGGGATTAATCTTATGATGAGGCGATTTAAGATGGTTCAAGAAACGATAAACTGGT 876
QY 816 ATCAGAGTATCTCCAAATGGAAGAAAGCAAGAAAGCAAGAAAGTCTGTTGTATGACCA 875
DB 877 CCAGAGATTAAGAAATTTGATGATTTGCGATATGCCAGAAACCTGAAATATATACAG 936
QY 876 GCTTAATAGTCCAAATTTTACGTTATGAGGAAT 911
DB 937 ACTGCATCATGCAAAATTCATGATTTATGATGACT 972

```

RESULT 16

```

US-10-835-516-61
; Sequence 61, Application US/10835516
; Publication No. US20040194162A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehnle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; TITLE OF INVENTION: create novel traits in transgenic organisms
; FILE REFERENCE: KAS-103XC1
; CURRENT APPLICATION NUMBER: US/10/835,516
; CURRENT FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/918,740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patencin version 3.0
; SEQ ID NO 61
; LENGTH: 7681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Operon C containing A. thaliana, S. cerevisiae, and Streptomyces
; OTHER INFORMATION: CL190
; OTHER INFORMATION: DNA, and R. capsulatus DNA
US-10-835-516-61

```

```

Query Match      10.2%; Score 132; DB 18; Length 7681;
Best Local Similarity 50.5%; Pred. No. 1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

```

```

QY 3 GTCMAAAGATTGACGACCTGGAAGAAAGCAATTTCTGCTGGTGAATTTGGTTCTTGA 62
DB 43 GTTGAGAGCTTCAGGCCCCAGGGAACGGTTACGTGGTGGATTTTACTTTTAA 102
QY 63 GCCAATTTATGATGCTTATGTGACAGCATTTGTCATGCAAGATGATGATTAACACC 122
DB 103 TACMAAATATGAGCAATTTGTAGTCGATTTATGCGAATATGCAATGCTGTAGCCATCC 162
QY 123 AAAAGAAC-----CAGTTTGAAGATCTAGAAATCAAAATTTCTTCAACCCCAATTGC 176
DB 163 TTACGGTTCATTCAGCAAGGCTGTGATTAAGTTGAAGCGGTGAAAGTAACAAATTTAA 222
QY 177 AAACGAGAAATGGGAATATCATCATCAATCAATACAG--AGAACCCAGAGAAATGCA 233
DB 223 AATATGGGAGTGGCTGTACCATATTAAGTCTTAAAGTGGCTTCATTCCTGTTTCATGAG 282
QY 234 GTCAGCATTAATTCATTTTGAAGGCACTATATTATCATGTTTGAAGTTATTCAC 293
DB 283 CGGATTAAGAACCTTTTCAATGAAAAGTTATCGCTAACGATTTAGCTACTTTAAAC 342
QY 294 GACCGAAGCATTTGATCTTGAA-----TCATCATTTTACTCAGACCCTGGAATATCA 344

```



```

; CURRENT APPLICATION NUMBER: US/10/835,516
; CURRENT FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/918,740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 58
; LENGTH: 7693
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Operon containing A. thaliana and S. cerevisiae DNA
US-10-835-516-58

Query Match      10.2%; Score 132; DB 18; Length 7693;
Best Local Similarity 50.5%; Pred. No. 1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGCTGTGTGATTTATTTGGTCTTGA 62
DB 3796 GTTGAAGACCTTCAGTGGCCCGAGAAAGCGTTACTAGCTGGTGGATTTATTTAGTTTGA 3855
QY 63 GCCAATTTATGATCTTATGTGACAGCATTTGCATCAGAAATGCAGTTATTAACACC 122
DB 3856 TACAAATATGAAGCATTTGTAGTCGATTCGGAGAAATGCATGCTGATCCCATCC 3915
QY 123 AAAAGAAC-----CAGTTGAAAAGATCTAGATCAAAATTTCTTCAACCCCATTTGC 176
DB 3916 TTACGGTTATGTCAGAGGGCTGATAGTTGAAGCGCTGTAAAGTAAACAAATTTAA 3975
QY 177 AAAAGGAGATGGGATATCATCATCATCATCAATTAACAG--AGAGCCCGAGAAATGTC 233
DB 3976 AGATGGAGTGGCTGCTACATTAAGTCTTAAAGTGGCTTATTCCTGTTTCATAGG 4035
QY 234 GTACGCGATTAATTCATTTTATGAGCAACTATATTCATGTTTATGCTTATTAACACC 293
DB 4036 CGGATCTAAGAACCTTTTCATTTGAAAAGTTATCGTAAAGTATTAAGTACTTTAAAC 4095
QY 294 GACCGAAGCATTTGATCTTGA---TCATCATTTACTCAGACCTTGATATCA 344
DB 4096 TAACATGAGCACTACTGCAATGAAACCTTGTCGTTATGATATTTTCTGTATGATGC 4155
QY 345 TTCACAGAAAGATCTAGAAACCAAGCATCTCGAATGAGAAAGAAATTTCTTAAACA 404
DB 4156 CTACCACTTCTCAGAGAGATGCTTACCGAATCGTGGCAAGAAAGATTGAGTTTCA 4215
QY 405 TTCTCGTGCATTTACCGAAGTGAAGAAAGACCGGATTAGTTTCATCGGACAGATTAGTGC 464
DB 4216 TTGCGACAGATTTGAAGAGATTTCCCAAAACAGGGCTGGGCTCTCGGACAGTTTAACTGC 4275
QY 465 AGTTGTTGCCAAGTTTATTTATCCATTTATCC-----CAATGTTATCAGTAC 515
DB 4276 AGTTTAACTACAGCTTTGGCTCTTTTATTTGATGGAACCTGAAATATATGATAGCA 4335
QY 516 GAATTAAGATTTTTCAGCAAGCTTGCACAGATTGCACATTTGTTATGCCCAAAAAGAT 575
DB 4336 ATATAGAGAAATTTATCATATTTAGCAAGATTTGCTCATTTGCAAGCTCAGGGTAAAT 4395
QY 576 AGGATCTGGGTTGATGTTGCACTGCAATTTATGCTGATTTATATGAAAGATTTC 635
DB 4396 TGGAAACGGGTTGATGTAAGCGGCGGAGATATGATCTATCAGATATAGAAAGATTCCC 4455
QY 636 GCCAGCTTTGATTAATGACGCTTTTCAAGTTTGAAGATGATCTGAGAAAGTTCCCCAC 695
DB 4456 ACCCGCATTAATCTTAATTTGCCAGATAT-----GGAAGTCTACTTACCGCAGTAA 4509
QY 696 AGAGTTGAAAAAATTTGATGAAGTAATCTGGGAATTCAAACATGAAGATGATCAATTAC 755
DB 4510 ACTGGCGCATTTGGTGTGATGAAGAAAGATGGAATATTTAGCATTAAGATCAACATTTAC 4569
QY 756 ATACGAATCAAGTTTATTAATGAGTGAAGTCAAGGGTGGCTCAGAAACCCAAATTTGAT 815

```

```

DB 4570 TTGCGGATTAATCTTATGATGAGCGCATATTAAGATGTTGTCAGAAACGTAAACCTGCT 4629
QY 816 ATACAGATTAATCTTATGATGAGCGCATATTAAGATGTTGTCAGAAACGTAAACCTGCT 875
DB 4630 CCAGAAAGTAAATAATTTGATGATTTGCGATATGCGATATGCGAAGAGCTTGAAATATATACAG 4689
QY 876 GCTTAATAGTGCATTTTACAGTTTATGAAAGAAATT 911
DB 4690 ACTGCATCATGCAAAATTTTATGATGATGAGTACT 4725

RESULT 19
US-09-918-740-59
; Sequence 59, Application US/09918740
; Publication No. US20030033626A1
; GENERAL INFORMATION:
; APPLICANT: Hann, Frederick
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XC1
; CURRENT APPLICATION NUMBER: US/09/918,740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 59
; LENGTH: 7695
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Operon B containing A. thaliana and S. cerevisiae DNA
US-09-918-740-59

Query Match      10.2%; Score 132; DB 10; Length 7695;
Best Local Similarity 50.5%; Pred. No. 1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGCTGTGTGATTTATTTGGTCTTGA 62
DB 28 GTTGAAGACCTTCAGTGGCCCGAGAAAGCGTTACTAGCTGGTGGATATTTAGTTTGA 87
QY 63 GCCAATTTATGATCTTATGTGACAGCATTTGCATCAGAAATGCAGTTATTAACACC 122
DB 88 TACAAATATGAAGCATTTGTAGTGGATTAATGCGAATGCAAGATGCTGTACCCATCC 147
QY 123 AAAAGAAC-----CAGTTGAAAAGATCTAGATCAAAATTTCTTCAACCCCATTTGC 176
DB 148 TTACGTTTATGCAAGGGCTGATTAAGTTGAAGTGGCTGTGAAGAAAGTAAACATTTAA 207
QY 177 AAAAGGAGATGGGAATATCATCATCATCAATTAACAG--AGAGCCCGAGAAAGTTCA 233
DB 208 AGATGGGAGTGGCTGCTACATTAAGTCTTAAAGTGGCTTCATTCCTGTTTCGATAGG 267
QY 234 GTACGCGATTAATTCATTTTATGAGCAACTATATTCATGCTTTATGCTTATTAACACC 293
DB 268 CGGATCTAAGAACCTTTTCATTTGAAAAGTTATCGTAAAGTATTAAGTACTTTAAAC 327
QY 294 GACCGAAGATTTGATCTTGA---TCATCATTTACTCAGACCTCGATATCA 344
DB 328 TAACATGAGCACTACTGCAATTAAGAACTTTGCTGTTATTTGATTTATTTCTGTATGATGC 387
QY 345 TTCACAGAAAGATCTGAACCAAGCATCTCGAATGAGAAAGAAATTTCTTAAACA 404
DB 388 CTACCACTTCTCAGAGAGATGCTTACCGAATCGTGGCAAGAAAGTTGAGTTTCA 447
QY 405 TTCTCGTGCATTTACCGAAGTGAAGAAAGACCGGATTAGTTTATGCGGACAGATTAGTGC 464
DB 448 TTGCGACAGATTTGAAGAGTTTCCCAAAACAGGGCTGGGCTCTCGGACAGTTTAACTGC 507
QY 465 AGTTGTTGCCAAGTTTATTTATCCATTTATCC-----CAATGTTATCAGTAC 515

```



```
Db 508 AGTTTAACTACACCTTGGCCCTCTTTTGTATCGACCTGGAAATATATAGACAA 567
Qy 516 GAATTAAGATATTTTGACACAGTTGACAGATTTGCACATTTGATATGCCAAAAAAGAT 575
Db 568 ATATAGAGAACTTATCTATTAATTTAGCAAGATTTGCTCATTTGCTAACCTCAGGGTAAAT 627
Qy 576 AGATCTGGTTTGTATGTTGCACTGCAATTTATGATCTGATTTATATAGAAATTTCA 635
Db 628 TGAAGCGGGTTTATGATGAGCGGCGACATATGATCTATCATGATATAGAAATTTCC 687
Qy 636 GCCAGCTTTGATTAATGACGTTTCAAGTTCTAGAAAGTATCTGAGAAAGTTCCAC 695
Db 688 ACCCGATTAATCTCTAATTTGCGAGATATT-----GGAAGTCTCTATTACGCGAGTAA 741
Qy 696 AGAGTTGAAAAAATTTGATTTGAAAGTAACTGGGAATTCAAACATGAAAGATATACCT 755
Db 742 ACTGGCCCATTTGGTTATATGAAAGACTGGAATTTATGATTAAGAAATTAACCATTTACC 801
Qy 756 ATACGGAATCAAGTTATTAATGGGTGACGTCAAGGGTGGCTCAGAAAACCCAAATTTGT 815
Db 802 TTGGGATTTAACTTTATGATGAGCGCATTTTAAGAAATGTTCAGAAACAGTAAATCGGT 861
Qy 816 ATACGAGTACTCCAAATGAAAAAGAAAGCCAGAAAGACTCTGTTGTATGACCA 875
Db 862 CCGAAGGTAAAAAATTTGATGATTCGATATGCCAGAAAGCTTGAAAAATATATACAG 921
Qy 876 GCTTAATAGTCCAAATTTACGTTTAAGAGAAAT 911
Db 922 ACTGATCATGCAAAATTTCTAGATTTATGATGAGACT 957

RESULT 20
US-10-835-516-59
; Sequence 59, Application US/10835516
; Publication No. US20040194162A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehnle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XCI
; CURRENT APPLICATION NUMBER: US/10/835,516
; CURRENT FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/918,740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 59
; LENGTH: 7695
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Operon B containing A. thaliana and S. cerevisiae DNA
US-10-835-516-59

Query Match 10.2%; Score 132; DB 18; Length 7695;
Best Local Similarity 50.5%; Pred. No. 1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
```

```
Qy 177 AAACGGAATGGGAATATCATATCATCAATACAG---AGAAACCCAGAAAGTTCA 233
Db 208 AGATGGGAGTGGCTGTACATATATAGTCTTAAAGTGGCTTCATTCCTGTTTCATAGG 267
Qy 234 GTTACGCAATTAATTCATTTTATGAGCAACATATTCATTCGTTTATAGTTATTAACCC 293
Db 268 CGATCTAAGAACCTTCATTTGATTAAGAAAGTTATCGTAAACGATATTTAGTACTTTAAC 327
Qy 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTTATCTGAGACCTGATATCA 344
Db 328 TAACATGACGACTACTGCAATAGAAACTGTGTTGCTATATGATATTTTCTGATGATGC 387
Qy 345 TTCAACAAGAAATCTGAAACCAACAGATCCCTCCGAAATGAGAAAAAATTTCTTTACCA 404
Db 388 CTACCATTTCTCGAAGATAGCGTTTACCGAATCTGTGCAACAGAAATGAGATTTTCA 447
Qy 405 TTCTGTGTCATTACCGAAGTGAAGAAGCCGATTTAGTTTATCGGAGATTTAGTGC 464
Db 448 TTGCAACAGAAATTTGAAGAAAGTTCCAAAACAGGGCTGCGGCTCTCGCAGGTTTATGAC 507
Qy 465 AGTTGTGCGACAAGTTTATATCCCATTTTATCCC-----CAATGTTATCAGTAC 515
Db 508 AGTTTAACTACAGCTTTGGCCCTCTTTTGTATCGACCTGGAATAATATATAGACAA 567
Qy 516 GAATTAAGTATTTTGGCAACAGTTGGACAGATTTGCAATTTGATTTGCCAAAAAAGAT 575
Db 568 ATATAGAGAAATTTATCTAATTTTATGACAAAGTGTCTATTTGCAAGCTCAGGGTAAAT 627
Qy 576 AGATCTGGGTTTGAATGTTGCAACTGCAATTTATGTCATGATGATATAGAAATTTCA 635
Db 628 TGAAGCGGGTTTGAATGATGAGCGGCGACATATGATCTATCATGATATTAAGATTTCC 687
Qy 636 GCCAGCTTTGATTAATGACGTTTCAAGTTCTAGAAAGTATCTGAGAAAGTTCCAC 695
Db 688 ACCCGATTAATCTCTAATTTGCGAGATATT-----GGAAGTCTCTATTACGCGAGTAA 741
Qy 696 AGAGTTGAAAAAATTTGATTTGAAAGTAACTGGGAATTCAAACATGAAAGATATACCT 755
Db 742 ACTGGCCCATTTGGTTATATGAAAGACTGGAATTTATGATTAAGAAATTAACCATTTACC 801
Qy 756 ATACGGAATCAAGTTATTAATGGGTGACGTCAAGGGTGGCTCAGAAAACCCAAATTTGT 815
Db 802 TTGGGATTTAACTTTATGATGAGCGCATTTTAAGAAATGTTTCAGAAACAGTAAATCGGT 861
Qy 816 ATACGAGTACTCCAAATGAAAAAGAAAGCCAGAAAGACTCTGTTGTATGACCA 875
Db 862 CCGAAGGTAAAAAATTTGATGATTCGATATGCCAGAAAGCTTGAAAAATATATACAG 921
Qy 876 GCTTAATAGTCCAAATTTACGTTTAAGAGAAAT 911
Db 922 ACTGATCATGCAAAATTTCTAGATTTATGATGAGACT 957

RESULT 21
US-09-918-740-62
; Sequence 62, Application US/09918740
; Publication No. US2003003626A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehnle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XCI
; CURRENT APPLICATION NUMBER: US/09/918,740
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 62
; LENGTH: 8224
; TYPE: DNA
```


ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Operon E containing A. thaliana, S. cerevisiae, Steptomyces sp C
OTHER INFORMATION: DNA,
OTHER INFORMATION: and R. capsulatus
US-09-918-740-62

Query Match 10.2%; Score 132; DB 10; Length 8224;
Best Local Similarity 50.5%; Pred. No. 1.1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

3 GTCAAAAGCATTTAGGACCTTGAAAAGCATTTCTTCTGCTGGGATTTGGTTCTTGA 62
43 GTTGAGAGCCTTACGTCGCCAGGGAAGGCTTACCTAGCTGGGATTTAGTTTGA 102
63 GCCAATTTATGATGCTATGTCAGCATGTCATGCAAGATGATGATGATTAACCC 122
103 TACAAATATGAAAGCATTTTGTAGTGGATTTAGGCAAGATGATGATGATTAACCC 162
123 AAAAGGAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCAACCCATTTGC 176
163 TTACGGTTATGCAAGGCTGATTAAGTTGAAGTCGCTGAAAGTAAACATTTAA 222
177 AAACGAGATGGGAATATCATATCATCAATATACG--AGAACCCAGAGATTTCA 233
223 AGATGGGAGTGGCTGATACCATATTAAGTCTTAAAGGCTTCACTTCTTTCATAGG 282
234 GTACACATTAATTCATTTTATGAGGCACTATATTCATGCTTTTATGCTTATATTCACC 293
283 CGATCTTAAGAACCTTCTTATTAATAAGTTATGCTTACGATTTAGCTATTTAAAC 342
294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTACTCAGACCTGGATATCA 344
343 TAAATGAGAGCATCTACGAATTAAGAACTTGTGTTATGATTTTCTCTGATGATGC 402
345 TTCAACAAGATATCTGAAACCAAGACATCTCGAATGAGAAAACATTTCTTTACCA 404
403 CTACCATTTCTCAGAGGATAGCGTACCAACATCGGCAACAGAAATTTGAGTTTCA 462
405 TTCTCGGCATTAACGAAGTGAAGAAAGCCGATTAAGTTTCACTCGGCAAGATTAAGTGC 464
463 TTGCAACAGATTAAGAAAGTTCCCAAAAGGCTGGGCTCTCGGCAAGTTAGTAC 522
465 AGTTGTTGCCACAAGTTTATATCCATTTTATCCC-----CAATGTTACAGTAC 515
523 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATGCGACCTGGAATTAATGAGCA 582
516 GAATTAAGATATTTTGACAACGTTGCAAGATTTGCACTTTGTTAGCCCAAAAAGAT 575
583 ATATAGAGAATTAATCATTAATTAGCAAGTTGCTCATTTGTCAGCTCAGGGTAAAT 642
576 AGATCTGGGTTGATGTTGTCACCTGCAATTTATGCTGATTAATTAAGATTTCA 635
643 TGAAGCGGGTTGATGTCAGCGGCGGAGCATATGATCTATCAGATTAAGAAATTTCC 702
636 GCCAGCTTGAATAAGACGTTTCAAGTTCTAGAAAGTATCTGAGAAAGTTCCAC 695
703 ACCCGATTAATCTTAATTTGCCAGATTT-----GGAAGTCTACTTAACGAGTAA 756
696 AGAGTTGAAAAAATGATTAAGAAATTAAGTGAATTTCAACATGAAGATGATCAATTAC 755
757 ACTGGGCAATTTGTTGATGTAAGAAAGTGAATTAATTAAGATTAACATTTTAC 816
756 ATACGGAATCAATTTATTAATGCTGACGTCAGAGGCTGTCAGAAACACCAATTTGT 815
817 TTGGGATTTAATCTTAATGATGCGCATTAATTAAGAAAGTTTCAAGAAACATTAACGCT 876
816 ATCAGAGTACTCAATGAAAAAGAAAGCAAGCAAGCAAGCAAGCTGTTGTTATGACCA 875
877 CCAAGAGTAAAAATTTGATGATTTGCGCATATGCGCAAGCAAGCTTGAATTAATTAACA 936
876 GCTTAATAGTCCCAATTTACAGTTTATGAAGAAAT 911

Db 937 ACTGATCATGCAAAATTTAGATTTATGATGATGACT 972

RESULT 22

US-10-835-516-62
Sequence 62, Application US/10835516
Publication No. US20040194162A1
GENERAL INFORMATION:
APPLICANT: Hahn, Frederick
APPLICANT: Kuehle, Adelheid
TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
TITLE OF INVENTION: create novel traits in transgenic organisms
FILE REFERENCE: KAS-103XCI
CURRENT APPLICATION NUMBER: US/10/835,516
PRIOR FILING DATE: 2004-04-28
PRIOR APPLICATION NUMBER: US/09/918,740
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/221,703
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn version 3.0
SEQ ID NO 62
LENGTH: 8224
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Operon E containing A. thaliana, S. cerevisiae, Steptomyces sp C
OTHER INFORMATION: DNA,
OTHER INFORMATION: and R. capsulatus
US-10-835-516-62

Query Match 10.2%; Score 132; DB 18; Length 8224;
Best Local Similarity 50.5%; Pred. No. 1.1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

3 GTCAAAAGCATTTAGGACCTTGAAAAGCATTTCTTCTGCTGGGATTTGGTTCTTGA 62
43 GTTGAGAGCCTTACGTCGCCAGGGAAGGCTTACCTAGCTGGGATTTAGTTTGA 102
63 GCCAATTTATGATGCTATGTCAGCATGTCATGCAAGATGATGATGATTAACCC 122
103 TACAAATATGAAAGCATTTTGTAGTGGATTTAGGCAAGATGATGATGATTAACCC 162
123 AAAAGGAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCAACCCATTTGC 176
163 TTACGGTTATGCAAGGCTGATTAAGTTGAAGTCGCTGAAAGTAAACATTTAA 222
177 AAACGAGATGGGAATATCATATCATCAATATACG--AGAACCCAGAGATTTCA 233
223 AGATGGGAGTGGCTGATACCATTAAGTCTTAAAGGCTTCACTTCTGTTGATAGG 282
234 GTACACATTAATTCATTTTATGAGCACTATATTCATGCTTTTATGCTTATTAAC 293
283 CGATCTTAAGAACCTTTTCAATTAATAAGTTATCGCTTAAGATTTAGCTTTAAAC 342
294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTACTCAGACCTGGATATCA 344
343 TAAATGAGAGCATCTACGAATTAAGAACTTGTGTTATGATTTTCTCTGATGATGC 402
345 TTCAACAAGATATCTGAAACCAAGACATCTCGAATGAGAAAAACATTTCTTTACCA 404
403 CTACCATTTCTCAGAGGATAGCGTTACCAACATCGGCAACAGAAATTTGAGTTTCA 462
405 TTCTCGGCATTAACGAAGTGAAGAAAGCCGATTAAGTTTCACTCGGCAAGATTAAGTGC 464
463 TTGCAACAGATTAAGAAAGTTCCCAAAAGGCTGGGCTCTCGGCAAGTTAGTAC 522
465 AGTTGTTGCCACAAGTTTATATCCATTTTATCCC-----CAATGTTACAGTAC 515
523 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATGCGACCTGGAATTAATGAGCA 582
516 GAATTAAGATATTTTGACAACGTTGCAAGATTTGCACTTTGTTAGCCCAAAAAGAT 575

Db 583 ATATAGAGAACTTATTCATATTTAGACAAAGTTGCTCATTTGTCAGAGCTCAGGTAAT 642
QY 576 AGGATCTGGGTTGATGTTGCACTGCAATTTATGTCGATTGATATATAGAAATTTCA 635
Db 643 TGAAGAGGGGTTGATGATAGCGGAGAGATATGATCTATGATATATAGAAATTTCC 702
QY 636 GCCAGCTTTGATTAATGACGTTTTCAGGTTCTAGAAAGTATCTGAGAGTTCCCCAC 695
Db 703 ACCCGATTAATCTCTAATTTGTCAGATATTT-----GGAAAGTCTACTTACCGCAGTAA 756
QY 696 AGAGTTGAAAAAATGATGTAAGTAACTGGAATTCAAACATGAAGAATGATACATTTAC 755
Db 757 ACTGGGCGCATTTGGTGTGATGAAGAAAGCTGAAATTTAGCATTTAAAGTAAACATTTAC 816
QY 756 ATACGGAATCAAGTTATTTAATGAGGTGACGTCAAGGGTGCTCAGAAACCCAAATTTGT 815
Db 817 TTGGGATTAATCTTTATGATGAGCGCATTTAAGAAATGCTTCAGAAACAGTAAACCTGT 876
QY 816 ATACAGATCTCCATGGAAGAAAGCAAGAAAGCTCTGTTGTATGACCA 875
Db 877 CCAGAAAGTAAAAAATTTGATATTTGCAATATGCCAAGAAAGCTTGAATAATATATACAG 936
QY 876 GCTTAATAGTGCATTTAAGTTAAGTAAGAAAT 911
Db 937 ACTGATCATGCAATTTAGATTTATGATGAGACT 972

RESULT 23

US-09-918-740-60
; Sequence 60, Application US/09918740
; Publication No. US20030033626A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehnle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XC1
; CURRENT APPLICATION NUMBER: US/09/918, 740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221, 703
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 60
; LENGTH: 8235
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Operon C containing A. thaliana, S. cerevisiae, and R. caosulatus
US-09-918-740-60

Query Match 10.2%; Score 132; DB 10; Length 8235;
Best Local Similarity 50.5%; Pred. No. 1,1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGACCTTGAAAGCATTTCTGCTGTGATATTTGGTTCTGA 62
Db 28 GTTGAAGGCTTCAGTGCCCGAGGGAAGCGTTACTAGCTGGGATATTTAGTTTAA 87
QY 63 GCCAATTTATGATCTATGTGACAGCATTTGCATACCAAGGCAAGCAAGTTATACAC 122
Db 88 TACAAATATGAAAGCATTTGTAGTCCGATTTATCGGAAGAAAGCAATGCTGTAGCCCATC 147
QY 123 AAAAGAAC-----CAGTTTGAAGAAATCTAGAAATCAAAATTTCTACCCCAATTTGC 176
Db 148 TTACGGTTTCATTCGACAGGCTGTGATTAAGTTGAAGCGGTGTAAGAAAGTAAACATTTTAA 207
QY 177 AAAAGGAGATGGGAATATCATATCATCAATCAATACAG--AGAACCCAGAGAAAGTTCA 233
Db 208 AGATGGGAGTGCTGTACCATATTAAGTCTTAAAGTGCTTCAATCTGTTTCATAGG 267
QY 234 GTACAGCATTAATTCATTTTATAGGCAACTATATTCATCTGTTTACGTTATATTTCAACC 293

Db 268 CCGATCTAAGAAACCTTTTCATGAAAAAGTAAATCGTAAAGTATTTAGCTATTAAACC 327
QY 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTTCTCAGACCCTGATATCA 344
Db 328 TAAATAGACGACTACTGCAATAGAAACTGTTTCTGTTATGATATTTTCTGATGATGC 387
QY 345 TTCAAGAAAGTACTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTTACCA 404
Db 388 CTACCATTTCTCAGAGATAGGCTTACCGAACATCGTGCAACAGAAATGATGATTTTCA 447
QY 405 TTCTGTCGCATTAACCGAATGAAAAAGCCGATTTAGTTTCAATGGCGAGATTTGCTGC 464
Db 448 TTGCGACAGAAATGAAGAAATTCGAAAAACAGGGCTGGGCTCTCGCGAGTTTATGTCAC 507
QY 465 AGTTGTGCGCAAGATTATATATCCCATTTTATCC-----CAATGTTATCAGTAC 515
Db 508 AGTTTAACTACAGCTTTTGCCCTCTTTTGTATGAGACCTGGAATAATATGATAGACA 567
QY 516 GAATAAAGTATTTTGCACACGTTTGACAGATTTGCAATTTGATATGCCCCAAAAAGAT 575
Db 568 ATATAGAGAAATTTATTCATTAATTTAGCAACAGTTGCTCATTTGCAAGCTCAGGTTAAAT 627
QY 576 AGGATCTGGGTTGATGTTGCAATCTGCAATTTATGCTGATTTATATAGAAATTTCA 635
Db 628 TGAAGCGGGTTGATGATGAGCGCGCGAGCATATGATCTATCAGATATAGAAATTTCC 687
QY 636 GCCAGCTTGAATAATGACGTTTCAAGTTCTAGAAAGTATCTGAGAAATTTCCCCAC 695
Db 688 ACCCGCATTAATCTCTAATTTGCCAGATATTT-----GGAATGCTACTTACGCGAGTAA 741
QY 696 AGAGTTGAAAAAATTTGATGAAAGTAACTGGAATTTCAACATGAAAGATTTACATTAC 755
Db 742 ACTGGGCGCATTTGGTGTATGAAAGAACTGGAATTTACATTAAGTAAACCATTTACC 801
QY 756 ATACGGAATCAAGTTATTAATGAGGTGACGTCAAGGGTGCTCAGAAACCCAAATTTGT 815
Db 802 TTGGGATTAATCTTATGATGAGCGCATTTATGAATGGTTTGAAGCAATTAACCTGCT 861
QY 816 ATACAGATCTCCATGGAAGAAAGCAAGAAAGCTCTGTTGTATGACCA 875
Db 862 CCAGAAAGTAAAAAATTTGATATGATTCGCATATGCAAGAAAGCTTGAATAATATATACAG 921
QY 876 GCTTAATAGTGCATTTAAGTTAAGTAAGAAAT 911
Db 922 ACTGATCATGCAATTTCTAGATTTATGATGAGACT 957

RESULT 24

US-10-835-516-60
; Sequence 60, Application US/10835516
; Publication No. US20040194162A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehnle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XC1
; CURRENT APPLICATION NUMBER: US/10/835, 516
; PRIOR FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/918, 740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221, 703
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 60
; LENGTH: 8235
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Operon C containing A. thaliana, S. cerevisiae, and R. caosulatus
US-10-835-516-60

Query Match 10.2%; Score 132; DB 18; Length 8235;
 Best Local Similarity 50.5%; Pred. No. 1,1e-22;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTTGAAAAGCATTTCTTGCTGCTGGATTTGGTTCTTGA 62
 DB 28 GTTGAAGAGCTTCAGTCCCGCAGGAAAGGCTTACTAGCTGGTGATTTAGTTTGA 87
 QY 63 GCCAATTTATGATGCTTATGTGACAGATGTCATCAGAAATGATGATGATTAACACC 122
 DB 88 TACAAATATGAGCATTTGTAGTGGATTTATGCGAAGATGATGCTTATGCCATCC 147
 QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCAACCCATTTGC 176
 DB 148 TTACGGTTCTTCAGAGGGCTGATTAAGTTGAAGTCGTGAAAGTAACATTTAA 207
 QY 177 AAACGAGATGGAATATCATATCATCAATACAG--AGAACCCAGAGAGTTCA 233
 DB 208 AGATGGGAGTGGCTGTACCATTAAGTCTTAAAGGCGCTTCAATTCCTGTTGATAGG 267
 QY 234 GTCAAGCATTAATTCATTTTAAAGGCACTATATTCATGCTTTAGCTTATATTCAC 293
 DB 268 CGGATTTAAGAACCTTTTCACTGAAAAGTTATGCTTACGATTTTACTTTAAAC 327
 QY 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTACTCAGACCCCTGATATCA 344
 DB 328 TAAATGAGAGCTACTGCAATAGAAATCTGTTCTGTTATGATTTTCTCTAGATGC 387
 QY 345 TTCAACAAGATATCTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTACCA 404
 DB 388 CTACCATTTCTCAGAGAGATGCGTTACCGAACATCGGCGAACAGAAATTTGAGTTTCA 447
 QY 405 TTCTCGGCATTAACGAGAGTGGAAAAAGCCGATTAAGTTTCAATCGGAGATTAAGTGC 464
 DB 448 TTGCGACAGATTAAGAAAGATTCCTCAAAACAGGCGTGGCTCTCGCAGAGTTAGTAC 507
 QY 465 AGTTGTGCCACAAGTTTATATCCATTTTATCCC-----CAATGTTACTAGTAC 515
 DB 508 AGTTTAACTACAGCTTTGGCTCTTTTGTATGCGACTGGAATAATATGAGCA 567
 QY 516 GAATTAAGATATTTTGCAACAGTTGCAAGATTTGCAATGTTTATGCCCCAAAAAAGAT 575
 DB 568 ATATAGAGAATTAATCATTAATTAGCAAGATTTGCTCATTTCAAGCTCAGGTAATAT 627
 QY 576 AGGATCTGGTTGATGTTGCACTGCAATTTATGCTGATGTATATGAAATTTCA 635
 DB 628 TGAAGCGGTTTGAATGAGCGGCGAGCATATGATCTATCAGATATGAAAGATTTCC 687
 QY 636 GCCAGCTTTGATTAATGACGTGTTCCAGTTCTAGAAAGTATCTGAGAAAGTTCCACC 695
 DB 688 ACCCGATTAATCTTAATTTGCCAGATATTT-----GGAAGTCTACTTACCGCATTA 741
 QY 696 AGAGTTGAAAAAATGATTAAGTAAGTAACGGAATTTCAACATGAAAGATGATCAATTAC 755
 DB 742 ACTGGGCGCATTTGATGATGAAGACATGAAATATTAACATTAAGTAACATTTAC 801
 QY 756 ATACGGAATCAATTTTAAATGCGTACGTCACAGGCTGCTCAAAAACCCCAATTTGT 815
 DB 802 TTGCGATTTAACTTTATGATGAGCGCATTAATTAAGAAATGTTTCAGAAACAGTAAATCGT 861
 QY 816 ATCAGAGTACTCCTCAATGGAAGAAAGCAAGAAAGCAAGAAAGCTCTGTTGATGACCA 875
 DB 862 CCAGAAAGTAAAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 921
 QY 876 GCTTAATAGTCCCAATTTTACAGTTTATGAAAGAAATTT 911
 DB 922 ACTGATCATGCAAAATTTCTAGATTTATGATGAGTACT 957

RESULT 25
 US-09-918-740-64
 ; Sequence 64, Application US/09918740
 ; Publication No. US20030033626A1

GENERAL INFORMATION:
 APPLICANT: Hahn, Frederick
 APPLICANT: Kuehne, Adelheid
 TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
 TITLE OF INVENTION: Create novel traits in transgenic organisms
 FILE REFERENCE: KAS-103X1
 CURRENT APPLICATION NUMBER: US/09/918,740
 PRIOR FILING DATE: 2001-07-31
 PRIOR APPLICATION NUMBER: 60/221,703
 PRIOR FILING DATE: 2000-07-31
 NUMBER OF SEQ ID NOS: 76
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 64
 LENGTH: 8400
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATUERS:
 OTHER INFORMATION: Operon G containing A. thaliana, S. cerevisiae, and S. pombe DNA
 US-09-918-740-64

Query Match 10.2%; Score 132; DB 10; Length 8400;
 Best Local Similarity 50.5%; Pred. No. 1,1e-22;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY 3 GTCAAAAGCATTTAGTGCACCTTGAAAAGCATTTCTTGCTGCTGGATTTGGTTCTTGA 62
 DB 28 GTTGAAGAGCTTCAGTCCCGCAGGAAAGGCTTACTAGCTGGTGATTTAGTTTGA 87
 QY 63 GCCAATTTATGATGCTTATGTGACAGATGTCATCAGAAATGATGATGATTAACACC 122
 DB 88 TACAAATATGAGCATTTGTAGTGGATTTATGCGAAGATGATGCTTATGCCATCC 147
 QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCAACCCATTTGC 176
 DB 148 TTACGGTTCTTCAGAGGGCTGATTAAGTTGAAGTCGTGAAAGTAACATTTAA 207
 QY 177 AAACGAGATGGAATATCATATCATCAAAATACAG--AGAACCCAGAGAGTTCA 233
 DB 208 AGATGGGAGTGGCTGTACCATTAAGTCTTAAAGGCTTCAATTCCTGTTGATAGG 267
 QY 234 GTCAAGCATTAATTCATTTTAAAGGCACTATATTCATGCTTTTATGCTTATATTCAC 293
 DB 268 CGGATTTAAGAACCTTTTCACTGAAAAGTTATGCTTACGATTTTATGATTAATTAAC 327
 QY 294 GACCGAAGCATTTGATCTTGAAA-----TCATCATTTACTCAGACCCCTGATATCA 344
 DB 328 TAAATGAGAGCTACTGCAATAGAAATCTGTTCTGTTATGATTTTCTCTAGATGC 387
 QY 345 TTCAACAAGATATCTGAAACCAAGACATCTCGAATGAGAAAAAATTTCTTTACCA 404
 DB 388 CTACCATTTCTCAGAGAGATGCGTTACCGAACATCGTGGCAAGAAATTTGAGTTTCA 447
 QY 405 TTCTCGGCATTAACGAGAGTGGAAAAAGCCGATTAAGTTTCAATCGGAGATTAAGTGC 464
 DB 448 TTGCGACAGATTAAGAAAGATTCCTCAAAACAGGCGTGGCTCTCGCAGAGTTAGTAC 507
 QY 465 AGTTGTGCCACAAGTTTATATCCATTTTATCCC-----CAATGTTACTAGTAC 515
 DB 508 AGTTTAACTACAGCTTTGGCTCTTTTGTATGCGACTGGAATAATATGAGCA 567
 QY 516 GAATTAAGATATTTTGCAACAGTTGCAAGATTTGCAATGTTTATGCCCCAAAAAAGAT 575
 DB 568 ATATAGAGAATTAATCATTAATTAGCAAGATTTGCTCATTTGCAAGCTCAGGTAATAT 627
 QY 576 AGGATCTGGTTGATGTTGCACTGCAATTTATGCTGATGTATATGAAAGATTTCA 635
 DB 628 TGAAGCGGTTTGAATGAGCGGCGAGCATATGATCTATCAGATATGAAAGATTTCC 687
 QY 636 GCCAGCTTTGATTAATGACGTGTTCCAGTTCTAGAAAGTATCTGAGAAAGTTCCACC 695
 DB 688 ACCCGATTAATCTTAATTTGCCAGATATTT-----GGAAGTCTACTTACCGCATTA 741
 QY 696 AGAGTTGAAAAAATGATTAAGTAAGTAACGGAATTTCAACATGAAAGATGATCAATTAC 755

```
Db      742 ACTGGGCAATTTGGTTGATGAGAAAGACTGGAAATATTAGATTAAAGTAACCAATTAC 801
Qy      756 ATACGGAATCAAGTTATTTATGAGTGAAGGCTGAGAAACCCCAATTGGT 815
Db      802 TTCGGGATTAATCTTATGATGAGTGGCAATATTAAGATGGTTGAAACAGTAACCTGGT 861
Qy      816 ATCAGAGTACTCCAAATGGAAGAAAGCCAGAAAGAAAGCTGTGTGTATGACCA 875
Db      862 CCAGAAAGTAAAGAAATTTGGATGATTCGCATATGCCAGAAAGCTTGAATAATATACAG 921
Qy      876 GCTTAATAGTCCCAATTTACAGTTTATGAAGAAAT 911
Db      922 ACTGCATCATGCAAAATTTAGATTATGATGAGACT 957

RESULT 26
US-10-835-516-64
; Sequence 64, Application US/10835516
; Publication No. US20040194162A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehnle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XC1
; CURRENT APPLICATION NUMBER: US/10/835,516
; CURRENT FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/918,740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 64
; LENGTH: 8400
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Operon G containing A. thaliana, S. cerevisiae, and S. pombe DNA
US-10-835-516-64

Query Match      10.2%; Score 132; DB 18; Length 8400;
Best Local Similarity 50.5%; Pred. No. 1.1e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

Qy      3  GTCAAAAGCATTTAGTGAACCTGGAAGCAATTTCTGTGCTGATTTGGTTCTTGA 62
Db      28  GTTGAGAGCTTCAGTGGCCCGGGAAGCGTTACTAGCTGTGGATTTTGTATTGA 87
Qy      63  GCCAATTTATGATGCTTATGTGACAGCATTTGCATCAGAAATGATGATTAACACC 122
Db      88  TACAATAATGAGCAATTTGTAGTGGATTTATCGGAAAGATGATGCTGTAGCCATGC 147
Qy      123  AAAAGAAC-----CAGTTTGAAGAATCTAGAAATCAAAATTTCTTCAACCCCAATTGC 176
Db      148  TTACGGTTTATTCGAAGGCTGATTAAGTTTGAAGTGGCTGGAAGTAACAAATTAA 207
Qy      177  AAACGGAGATGGGAATATCATCATCAATCAATACAG--AGAAGCCAGAGAAATTTCA 233
Db      208  AGATGGGGAGTGGCTGTACCATATTAAGTCTTAAAGTGGCTTCATTCCTGTTCGATAG 267
Qy      234  GTACAGCATAAATCCATTTTATGAGGCAACTATATTCATCGTTTATGCTTAATTCAC 293
Db      268  CGATCTAAGAACCTTTTATTTAAGAAAGTATCGTAAAGTATTTAGCTACTTTAAAC 327
Qy      294  GACCGAAGCATTTGATCTTGAAG-----TCATCATTACTCAGACCTGATATGCA 344
Db      328  TAAACATGACGACTACTGCAATGAGAACTTGTTCGTTATGATTAATTTCTGTGATG 387
Qy      345  TTCACAGAAAGATGATGAAACCAAGCATCTCGAATGAGAAAGAAATTTCTTTACCA 404
Db      388  CTACCATCTCTCAGAGGATAGCGTTTACGAAACATCGTGCAACAGAAATGATTTTGA 447
```

```
Qy      405  TTCCTGTCATTAACCGAAGTGAAGAACCGGATTTAGTTGATCGGACGATTAAGTGC 464
Db      448  TTCGACAGAAATTTGAAGAAATTTCCCAAAACAGGGCTGGGCTCTCGGACGATTAAGTGC 507
Qy      465  AGTTGTGCAACAATTTATTTATCCCATTTTATCCC-----CAATGTTATCAGTAC 515
Db      508  AGTTTAACTTACAGCTTGGCCCTCTTTTATTCGAGCTGGAAGAAATATGATACACA 567
Qy      516  GAATAAGATATTTTGCACACGTTGACAGATTCGACATTTGTTATGCCCAAAAAAGAT 575
Db      568  ATATAGAGAGATTTATTTCAATTTTACAGAAATTTGCTATTTTCAAGCTCAGGTTAAAT 627
Qy      576  AGAATCTGGGTTTATGTTGCAATGCAATTTATGCTGATTTGATATAGAAATTTCA 635
Db      628  TGAAGCGGTTTATGATGATGCGGCGAGCATATGATCTATCAGATATAGAAATTTCCC 687
Qy      636  GCCAGCTTTGATTAATAGAGCTGTTTACAGGTTTCAAGAAAGTATCCTGAGAAATTTCCCA 695
Db      688  ACCCGCATTTAATCTCTAATTTGCGAGATATTT-----GGAAGTCTACTTACGCGATGA 741
Qy      696  AGAGTTGAAAAAATTTGATTTGAAGTAACTGGGAATTTCAACATGAAGATGTACATTACC 755
Db      742  ACTGGGCAATTTGGTTGATGAGAAAGACTGGAATTTAGATTTAAAGTAACCAATTACC 801
Qy      756  ATACGGAATCAAGTTATTAATGAGTGAAGCTCAAGGCTGCTGAGAAACCCCAATTGGT 815
Db      802  TTCGGGATTAATCTTATGATGAGGCGATATTAAGATGGTTTCAAGAAACGTAACCTGGT 861
Qy      816  ATCAGAGTACTCCAAATGGAAGAAAGCCAGAAAGAAAGCTGTGTGTATGACCA 875
Db      862  CCAGAAAGTAAAGAAATTTGGTATGATTCGATATGCGACAGAAACCTGGAATAATATACAGA 921
Qy      876  GCTTAATAGTCCCAATTTACAGTTTATGAAGAAAT 911
Db      922  ACTGCATCATGCAAAATTTAGATTATGATGAGACT 957

RESULT 27
US-10-006-909-7
; Sequence 7, Application US/10006909
; Publication No. US20030148479A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: KIM, SEON-WON
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: YOSHIKUNI, YASUO
; APPLICANT: NEWMAN, JACK
; APPLICANT: KHEBNIKOV, ARTEM VALENTINOVICH
; TITLE OF INVENTION: BIOSYNTHESIS OF ISOBERENTYL PYROPHOSPHATE
; FILE REFERENCE: 2000-0007
; CURRENT APPLICATION NUMBER: US/10/006,909
; CURRENT FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 9253
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: "single operon" nucleotide sequence
US-10-006-909-7

Query Match      10.2%; Score 132; DB 15; Length 9253;
Best Local Similarity 50.5%; Pred. No. 1.2e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

Qy      3  GTCAAAAGCATTTAGTGAACCTGGAAGCAATTTCTGTGCTGATGATTTGGTTCTTGA 62
Db      6259  GTTGAGAGCTTCAGTGGCCCGGGAAGCGTTACTAGCTGTGTGATTTATTTAGTTTGA 6318
```

QY	63	GCACATTTATGATGGTTTATGACACGACATTTGTCATCACGAATGATGAGTTATTAACCC	122
Db	6319	TACAAAATATGACATTTGTATGTCGATTTATCGCAGAAATGATGCTGTAGCCCATCC	6378
QY	123	AAAGGAAC-----CAGTTGAAAGAACTAGAACTCAAAATTTCTTCAACCCCAATTTGC	176
Db	6379	TTACGGTTCAATTCGAAAGGGTCTGATTAAGTTGAAGTGGGTGAAAGATGAACAATTTAA	6438
QY	177	AAAGGAATATGGGAATATCACTATTCATCAAAATACAG--AGAAGCCAGAGAGTTCA	233
Db	6439	AGATGGGAGTGGCTGTACCATATTAAGTCTTAAAGTGGCTTCATTCCTGTTTCGATAGG	6498
QY	234	GTCAAGCAATTAATCCATTTTATGAGGCACTATATTCATCGTTTATGCTTATATTAACCC	293
Db	6499	CGGATCTTAAGAACCTTTTCATTGAAAAAGTTATCGCTTAACGATTTATAGCTACTTTAAACC	6558
QY	294	GACCGAGCAATTTGATCTTGAAA-----TCATCATTTACTCAGACCCCTGGATATCA	344
Db	6559	TAAATGAGACACTACTCAATATAGAACTTGTCGTTATATGATATTTTCTCTGATGATGC	6618
QY	345	TTCAACAAGAAATACTGAACCAAGACATCTCTCGAATGGAAGAAAAAATTTCTTTACCA	404
Db	6619	CTACCATTTCTGAGAGATACGTTTACCGAACATCTGTGGCAACAGAAATTAAGTTTCA	6678
QY	405	TTCTGTGTCACATTAACGAATGGAAGAAAGCCGATTTAGTTTCATCGGAGGATTAAGTGC	464
Db	6679	TTGCGACAGATTTGAAGAAAGTTCCAAACAGGGGTGGGCTCTCTGGCAGGTTTATGTCAC	6738
QY	465	AGTTGTGCCACAAGTTTATTAATCCATTTTATCCC-----CAATGTTATCAATAC	515
Db	6739	AGTTTAACTACAGCTTTGGCCCTCTTTTGTGATCGGACCTGGAAAAATATGTAACAA	6798
QY	516	GAATTAAGATATTTTGGCAACGTTGCAACATTTGCACATTTGATGCCCCAAAAAGAT	575
Db	6799	ATATAGAAAGTATTCATTAATTTAGACAAGTTGCTCATTTGTCAGCTCAGGGTAAAT	6858
QY	576	AGGATCTGGGTTTGATGTTGCAACGTCAATTTATGTCGATTTGATATTAAGAAATTTCA	635
Db	6859	TGGAAGCGGGTTTGATGTAGGGGGGGGACGATATGATCTATCAGATATTAAGATTTCC	6918
QY	636	GCCAGCTTGTATTAATGACGTGTTTCAGGTTCTAGAAAGTATCTCGAAGATTTCCAC	695
Db	6919	ACCCGCATTAATCTCTAATTTGCGCAGATATT-----GGAAGTGTACTTTCGCGCAGTAA	6972
QY	696	AGAGTGAAGAAAAATTGATTTGAAGAAATCGGGAATTCAAACATGAAGAAATGATACCTTACC	755
Db	6973	ACTGGCGCATTTGGTTGATGAAAGAACTGGAAATTTTCGATTAAGATTAACATTTACC	7032
QY	756	ATACGGAATCAAGTTAATTAATGGGTGACGTCAAGGGTGGCTCAGAAACACCAAAATGGT	815
Db	7033	TTCCGGATTAATCTTTATGATGGGGGATTTTAAGATGTTTCAGAAACGTAATAAATCTGT	7092
QY	816	ATCACAGATCTCCCATATGAAAAAGAAAGACCGAAGAAAGCTCTGTGTGATGACCA	875
Db	7093	CCAGAAGGTAAAAAATTTGATGATTTCCGATATGCGAAGAAAGCTTGAATAATATATACAGA	7152
QY	876	GCTTAATATGTCGAATTTACAGTTTATGAGGAATTT	911
Db	7153	ACTGATCATCAAAATTTCAAGATTTATGATGAGTACT	7188
RESULT 28			
US-10-411-066-7			
Sequence 7, Application US/10411066			
Publication No. US20040005678A1			
GENERAL INFORMATION:			
APPLICANT: KEASLING, JAY			
APPLICANT: MARTIN, VINCENT			
APPLICANT: PITERA, DOUGLAS			
APPLICANT: WITHERS III, SYDOR T.			
APPLICANT: NEWMAN JACK			
TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE			

	Query Match	10.24;	Score 132;	DB 16;	Length 9253;
Query	3	GTCAAAAGCATTTAGTGCACCTGGAAAAAGCATTTCTTGCTGGTGGATTTGGTTCTTGA	62		
Db	6259	GTTGAGAGCCTTCAGTGCCTCCAGGAAAGCGTTTCTGCTGGTGGATTTAGTTTGA	6318		
Query	63	GCCAAATTTAGTGCCTTATGTCAGAGATTGTCATCGATCGATGAGTTAATACCC	122		
Db	6319	TACAAAATATGAAACATTTGTATGTCGGATTTACGGCAAGATGATGATGAGCCCATCC	6378		
Query	123	AAAGAAGAC-----CAGTTTGAAGAAGATCAGATCAAAATTTCTTACCCCAATTTGC	176		
Db	6379	TTACGTTTCATGTGCAAGGCTGTGATTAAGTTGAAGTGCCTGTGAAGTAACATTTAA	6438		
Query	177	AAACGGAATGGAATATCATCATATCATCAATACAG--AGAGCCGAGAGAATTCA	233		
Db	6439	AGATGGGAGAGGCTGTACATATAGTCTTAAAGGAGCTTCATTCCTGTTTGCATAGG	6498		
Query	234	GTCACGATTAATTCATTTTAAAGGCACTATATTCATGTTTACGTATATTTCAACC	293		
Db	6499	CGAGTCTAAGAACCTTTTCATGAAAAAGTTATGCTAAGCTATTTAGCTATTTAAACC	6558		
Query	294	GACCGAAGCATTTTATCTTGA--TCATCATTTACTCAAGCCTGGATATCA	344		
Db	6559	TAAATGAGAGCATCTGCAATAGAAACCTGTTCTGTTATGATTTTCTCTGATGATGC	6618		
Query	345	TTCACAAGAAGATCTGAACCAAGACATCTCGAATGAGAAAAACATTTCTTTACCA	404		
Db	6619	CTACATCTCAGAGAGATAGCTTACGAACATCGTGGCAACAGAGATTTGATTTTCA	6678		
Query	405	TTCTCTGTCATTACCGAATGAAAAAGCCGATTAAGTTTCATGCGCAGATTAAGTCTC	464		
Db	6679	TTCCACAGAAATGAAAGATTTCCCAAAACAGGGCTGGGCTCTCCGCAAGTTTAACTGC	6738		
Query	465	AGTTGTTGCCAACAGTTTATTCATTCATTCATTCCTC-----CAATGTTATCAGTAC	515		
Db	6739	AGTTTAACTACAGCTTTGGCTCTCTTTTGTATCGAACCTGGAAAAATATATGAGACA	6798		
Query	516	GAATTAAGATATTTTGGACAACGTTTGCACAGATTGCACTTGTATGCCCCAAAAAGAT	575		
Db	6799	ATATAGAGAGTTATTTATATTTTAAAGCAAGTTGCTCATTTGTCAACCTCAGGTAATAT	6858		
Query	576	AGATCTGGGTTTATGTTGCAACTGCATTTATGTTCTGATGTATATAGAGATTTC	635		
Db	6859	TGGAAGCGGGTTGATGTAGCGCGGACGACATATGATCTATCAGATATAGAAATTTCC	6918		
Query	636	GCCAGCTTTATTAATACAGTGTTCAGTTTCTAAGAAAGTATCTCGAAGATTCCCCAC	695		
Db	6919	ACCCGCAATTAATCTTAATTTGCCAGATTT-----GGAAGTCTCATTTACGCAAGTAA	6972		
Query	696	AGAATTAATAAAATTTGATGAAGATTAATCGGGAATTTCAACATGAAAGATGTACATTACC	755		
Db	6973	ACTGGCGCATTTGGTTGATGAAGAGCTGGAAATTTATCGATTTAAAGTAAACCATTTACC	7032		
Query	756	ATACGGAATCAAGTTATTAATGGGTGACGTCAAGAGGTGGCTCAGAAAACCCAAATTGGT	815		

```

Db      7033 TTCCGAGATTAACTTATAGATGGCGATATTAGAAATGGTTCAGAAACGTAAACTGGT 7092
Qy      816 ATCAGAGACTCCATGGAAGAAAGCAAGAGAAAGCTGTGTGTATGACCA 875
Db      7093 CAGAAAGGTAAATAATGGTATGATTCGATATGCCAGAAAGCTTGAATAATATACAGA 7152
Qy      876 GCTTAATAGTCCAAATTACAGTTTATGAGGAATT 911
Db      7153 ACTGATCATGCATAATTCTAGATTATGATGATGACT 7188

RESULT 29
US-09-918-740-72
; Sequence 72, Application US/09918740
; Publication No. US20030033626A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XC1
; CURRENT APPLICATION NUMBER: US/09/918,740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 13917
; TYPE: DNA
; ORGANISM: Artificial Sequence
; NAME/KEY: misc.feature
; LOCATION: (1..7)
; OTHER INFORMATION: Placid transformation vector pHK04, containing Operon B, contain
US-09-918-740-72

Query Match      10.2%; Score 132, DB 10; Length 13917;
Best Local Similarity 50.5%; Pred. No. 1,4e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

Qy      3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGTCTGTGTATTTGGTTCTTGA 62
Db      4137 GTTGAGAGCCTTCAGTGCCTCCAGGAAAGCTTACTGCTGTGTATTTAGTTTGA 4196
Qy      63 GCCAATTATGATGCTTATGTGACAGCATTTGTATCAAGATGCATGCAATTATACACC 122
Db      4197 TACAAATATGAGCATTTGTAGTCGATTTATCGCAAGATGATGCTGTAGCCCATCC 4256
Qy      123 AAAAGAAC-----CAGTTGAAAAGATCTAGAAATTTCTTCACCCCAATTGC 176
Db      4257 TTACGGTTTCATTTGCAAGGGCTGTATTAAGTTGAAGTCGTGTGAAAAGTAAACATTTAA 4316
Qy      177 AAACGGAATGGAATATCATATCATCAATACAG---AGAAGCCAGAGAGTTCA 233
Db      4317 AGATGGGAGTGGCTGTACATATTAAGCTTAAAGGGCTTCACTTCTGTGTGATAGG 4376
Qy      234 GTCAAGCATTAATCCATTTTGAAGGCACTATATTCATGCTTTTATGCTTATATTCACC 293
Db      4377 CGGATCTTAAGAACCTTTTCATTGAAAAAGTTATCGTAAACGTAACTTATAGCTTCTTAAACC 4436
Qy      294 GACCGAAGCATTTGATCTTGA-----TCATCATTTACACAGCCCTGGATATCA 344
Db      4437 TAAATGAGCAGCATCTAGCAATTAAGAACTTGTGTATTAATTTTCTGTGATGATGC 4496
Qy      345 TTCAAGAAATATCTGAAACCAAGACATCTCGATGAGAGAAAAACATTTCTTTACA 404
Db      4497 CTACCATTTCTGAGAGGATAGGTTAACGAACATGTGGCAACAGAAATTTGAAGTTTCA 4556
Qy      405 TTCTGTGTCATTACCGAAGTGGAAAAAGCCGATTTAGCTTCATGCGAGATTAATGTC 464

```

```

Db      4557 TTCCACAGAAATTGAGAAAGTTCACAAAACAGGGCTGGGCTCTCGGCAAGTTTATGTCAC 4616
Qy      465 AGTTGTTGCCCAAGATTATATATCCATTTATATCC-----CATGTTATCAGTAC 515
Db      4617 AGTTTAACTACACTTGTGGCTCTTTTGTATTCGAGCTGTGAAAAATATATGACAA 4676
Qy      516 GAATTAAGATATTTTGCACAAAGTTGACAGATTGCAATTGTTATGCCCCAAAAAGAT 575
Db      4677 ATATAGAGAAATTTATTCATTAATTAGACAAAGTGTCTCATTTGCAAGCTCAGGGTAAAT 4736
Qy      576 AGGATCTGGTTTGTATGTTGCACTGCAATTTATGCTGTATGATTAATAGATTCA 635
Db      4737 TGGAAAGCCGGTTTGAATGAGGCGGCAATATGATCTATACGATATTAAGATTTCC 4796
Qy      636 GCCAGCTTGAATTAATGACGTGTTTCTAGAAAGTATCCTGAGAAATTTCCAC 695
Db      4797 ACCCGCATTAATCTTAATTTGCCAGATTT-----GGAAGTGTACTTACGGCAGTAA 4850
Qy      696 AGAGTTGAAAAAATGATGAAAAGTAACTGGGAATTCAAACATGAAAGTATACCTACC 755
Db      4851 ACTGCGCATTTGGTTGATGAAAGAACTGGAATTTACGATTTAAAGTAAACATTTACC 4910
Qy      756 ATACGAATCAAGTTATTAATGCGTGAAGTGAAGGCTGCAAGAACCCCAATTTGAT 815
Db      4911 TTCCGGAATTAATCTTATGATGAGCGCATTTTAAGATGCTTCAGAAACAGTAAACTGGT 4970
Qy      816 ATCAGAGTACTCCATGGAAGAAAGCAAGAAAGCTCTGTGTGTATGACCA 875
Db      4971 CAGAAAGGTAAATAATGGTATGATTCGATATGCGAAGAAAGCTTGAATAATATATACAGA 5030
Qy      876 GCTTAATAGTCCAAATTACAGTTTATGAGGAATT 911
Db      5031 ACTGATCATGCATAATTCTAGATTATGATGATGACT 5066

RESULT 30
US-10-835-516-72
; Sequence 72, Application US/10835516
; Publication No. US20040194162A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103XC1
; CURRENT APPLICATION NUMBER: US/10/835,516
; PRIOR FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/918,740
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 13917
; TYPE: DNA
; ORGANISM: Artificial Sequence
; NAME/KEY: misc.feature
; LOCATION: (1..7)
; OTHER INFORMATION: Placid transformation vector pHK04, containing Operon B, contain
US-10-835-516-72

Query Match      10.2%; Score 132, DB 18; Length 13917;
Best Local Similarity 50.5%; Pred. No. 1,4e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

Qy      3 GTCAAAAGCATTTAGTGCACCTGGAAAAGCATTTCTTGTCTGTGTATTTGGTTCTTGA 62
Db      4137 GTTGAGAGCCTTCAGTGCCTCCAGGAAAGCTTACTGCTGTGTATTTAGTTTGA 4196
Qy      63 GCCAATTATGATGCTTATGTGACAGCATTTGTATCAAGATGCATGCAATTATACACC 122

```

```

Db      4197 TACAAATATGAAGCAATTTGTAGTCGGATTAATCGGCAAGATGATCTGTAGCCATCC 4256
QY      123 AAAAGGAAAC-----CAGTTGAAAGAAATCTAGAAATTAATTTCTTACCCCAATTTGC 176
Db      4257 TTACGGTCTATTGCAAGGGTCTGTATAGTTGAAGTCGGTGAAGAAATGAACATTTAA 4316
QY      177 AAACGGAATGGAATATCATATCATCAATATACAG---AGAGCCAGAGAGTTCA 233
Db      4317 AGATGGGAGGTGGCTGTACATATAAGTCTTAAAGGGCTTCAATCTGTTCGATAGG 4376
QY      234 GTACGCATTAATTCATTTTAAAGGCACTATATTTATGTTTACCTTATATTCACC 293
Db      4377 CGGATCTAAGAACCTTTTCAATTGAAAAGTATCGCTAAGATTTTGAACCTTTAAACC 4436
QY      294 GACCGAAGCATTTGATCTTGA---TCATCATTTACTCAGACCCCTGATATCA 344
Db      4437 TAAATGAGAGCACTATCTGCATTAAGAACTTGTCTGATTAATTTTCTCTGTATGATGC 4496
QY      345 TTCACAGAAGATATCTGAAACCAAGATCTCGATGAGAGAAAACATTTCTTTACCA 404
Db      4497 CTACCATTTCTCAGAGAGATAGCGTTACGAAACATCGTGGCAACAGAGATTTGAGTTTCA 4556
QY      405 TTCTCGTCCATTACCGAATGGAAGAACCGGATTAGTTTCATCGCAGAGATTAGTGC 464
Db      4557 TTGCAACAGAAATTAAGAAAGTCCCAAAACAGGCGTGGCTCCCTCGCAGGTTTATGATC 4616
QY      465 AGTTGTTGCCAAGTTTATTTATCCATTTTATCCC-----CAATGTATCAGTAC 515
Db      4617 AGTTTAACTACAGCTTGGCTTCCTTTTGTATCGACCTGGAATAATATGACAA 4676
QY      516 GAATTAAGATATTTTGCAACGTTGCAAGATTGACATTTGATGCCCCAAAAGAT 575
Db      4677 ATATAGAGAAATTAATCATATTTAAGCAAGATTGCTCATTTGCAAGCTCAGGGTAAAT 4736
QY      576 AGGATCTGGTTGATGTTGCAACTGCAATTTATGCTGATGTATATGAAGATTCA 635
Db      4737 TGGAAACGGGTTGATGATGACGGCGGACGATGATGATCATGATATGAAGATTCCC 4796
QY      636 GCCAGCTTGAATGAAGGTGTTTCAAGTTCTAAGAAAGATCCGGAAGTTCCCCAC 695
Db      4797 ACCCGATTAAATCTTAATTTGCGCAAGTATTT-----GGAAGTCTACTTAACGAGTAA 4850
QY      696 AGAGTTGAAAAAATTTGATGAAGTAATCGGAATTAACAATGAAGATGATCAATACC 755
Db      4851 ACTGGGCACTTGTGTTGATGAAGAAAGCTGGAATTTACATTAAGTAACCATTTACC 4910
QY      756 ATACGAATCAAGTTATTAATGAGGTGACGTCAAGGGTGGCTCAGAAAACCCAAATTTGT 815
Db      4911 TTGGGATTAACTTTATGATGAGGCGCATTTAAGAAATGTTCAAGAAACAGTAAACTGGT 4970
QY      816 ATACAGATCTCCATATGAAAAAGAAAAAGCAAGAAAGATCTCTGTGTATGACCA 875
Db      4971 CCAGAAAGTAAAAAATTTGATGATGATGCGCATATGCGCAAAAGCTTGAATAATATACGA 5030
QY      876 GCTTAATAGTCCATTTACGTTTATGAAGAAAT 911
Db      5031 ACTGATCATGCAAAATTTCTAGATTTATGATGACT 5066

```

RESULT 31

```

US-09-918-740-74
; Sequence 74, Application US/09918740
; Publication No. US20030033626A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Frederick
; APPLICANT: Kuehle, Adelheid
; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
; FILE REFERENCE: KAS-103JC1
; CURRENT APPLICATION NUMBER: US/09/918,740
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/221,703

```

```

; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 14623
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1..7)
; OTHER INFORMATION: Plasmid transformation vector pHK08, containing Operon G, contai
US-09-918-740-74

Query Match
Best Local Similarity 50.5%; Pred. No. 1,5e-22;
Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;

QY      3 GTCAAAAGCATTTAGTGCACCTGAAAAGCAATTTCTGTGTGATATTTGTTCTTGA 62
Db      2621 GTTGAAGACCTTCAGTGCCTCCAGGAAAGGCTTACTAGTGTGATATTTAGTTTGA 2680
QY      63 GCCAATTTATGATGCTTATGTGACAGATTGTCAATCAAGATGATGCAATTAACACC 122
Db      2681 TACAAATATGAAGCAATTTGTAGTGGATTAATGCGAAGATCATGCTGTAGCCATCC 2740
QY      123 AAAAGGAAAC-----CAGTTGAAAGAAATCTAGAAATTAATTTCTTACCCCAATTTGC 176
Db      2741 TTACGGTCTATTGCAAGGGTCTGTATAGTTTGAAGTCGGTGAAGAAATGAACATTTAA 2800
QY      177 AAACGGAATGGAATATCATATCATCAATATACAG---AGAGCCAGAGAAAGTTCA 233
Db      2801 AGATGGGAGGTGGCTGTACATATAAGTCTTAAAGGGCTTCATTCCTGTTTCATAGG 2860
QY      234 GTACGCATTAATTCATTTTAAAGGCACTATATTCATGTTTATGCTTATATTCACC 293
Db      2861 CGGATCTAAGAACCTTTTCAATTGAAAAGTATCGCTAAGATTAAGTACTTAAACC 2920
QY      294 GACCGAAGCATTTGATCTTGA---TCATCATTTACTCAGACCCCTGATATCA 344
Db      2921 TAAATGAGAGCACTATCTGCATTAAGAACTTGTGTTGATTAATTTTCTCTGTATGATGC 2980
QY      345 TTCACAGAAGATATCTGAAACCAAGATCTCGAATGAGAAAACATTTCTTTACCA 404
Db      2981 CTACCATTTCTCAGAGAGATAGCGTTACCAACATCGTGGCAACAGAAAGATTGAGTTTCA 3040
QY      405 TTCTCGTCCATTACCGAAGTGAAGAACCGGATTAGTTTCAATCGCAGAGATTAGTGC 464
Db      3041 TTGCAACAGAAATTAAGAAAGTCCCAAAACAGGGCTGGGCTCCCTCGCAGGTTTATGATC 3100
QY      465 AGTTGTTGCCAAGTTTATTTATCCATTTTATCCC-----CAATGTATCAGTAC 515
Db      3101 AGTTTAACTACAGCTTGGCTTCCTTTTGTATGAGACCTGGAATAATATGACAA 3160
QY      516 GAATTAAGATATTTTGCAACGTTGCAAGATTGCAATTTGATGCCCCAAAAGAT 575
Db      3161 ATATAGAGAAATTAATCATATTTAGCAAGATTTGCTCATTTGCAAGCTCAGGGTAAAT 3220
QY      576 AGGATCTGGTTGATGTTGCAACTGCAATTTATGCTGATGTATATGAAGATTCA 635
Db      3221 TGGAAACGGGTTTGAATGAGCGGCGGACGATGATGATCATATGATATGAAGATTCCC 3280
QY      636 GCCAGCTTGAATGAAGACGTGTTTCAAGTTCTAAGAAAGATCCGGAAGATTTCCCCAC 695
Db      3281 ACCCGATTAAATCTTAATTTGCGAAGTATTT-----GGAAGTCTACTTAACGAGTAA 3334
QY      696 AGAGTTGAAAAAATTTGATGAAGTAATCGGAATTAACAATGAAGATGATCAATACC 755
Db      3335 ACTGGGCACTTGTGTTGATGAAGAAAGCTGGAATTTAGCATTAAGTAACATTTACC 3394
QY      756 ATACGAATCAAGTTATTAATGAGGTGACGTCAAGGGTGGCTCAGAAAACCCAAATTTGT 815
Db      3395 TTGGGATTAACTTTATGATGAGGCGATTTAAGAAATGTTTCAAGAAACAGTAAACTGGT 3454

```


QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCACCCCAATTGC 176
 Db 2741 TTACGGTTCATGTGCAAGGCTGTGATTAAGTTGAAGTGTGTAAGAAATTAACAATTAA 2800
 QY 177 AAACGGAAATGGGAATATCAATATCATCAATAATACAG--AGAACCCAGAGAATTCA 233
 Db 2801 AGATGGGAGATGGGTACATATTAAGTCTTAAAGTGGCTTCAATTCCTGTTTGCATAG 2860
 QY 234 GTACGGATTAATCCATTTTATAGAGCACTATATTCATGCTTTTACCTTATATCAACC 293
 Db 2861 CGGATCTTAAGAACCTTTCATGTAAGAAAGTTATGCTTAACGTATTTAGCTACTTTAAAC 2920
 QY 294 GACCGAAGCATTTGATCTTGAA-----TCATCATTTTCTCAGACCTCGATATCA 344
 Db 2921 TTAATGAGAGCACTACGCAATAGAAACTTGTGTTATGATATTTCTCTGATGATGC 2980
 QY 345 TTCAACAAGAAATATCTGAACCAAGACATCTCGAATGAGAGAAAAAATTTCTTTTACA 404
 Db 2981 CTACCAATTCAGAGGATAGCGTTACCGAACATCGTGGCAACAGAAATTAAGTTTCA 3040
 QY 405 TTCTCGGCATTAACCAAGTGAAGAACCGGATTAGTTTATCGGCAAGATTAGTGTG 464
 Db 3041 TTGCGACAGAATTGAAGAAAGTTCCCAAAACAGGGCTGGCTCTCGGCAAGTTTACAC 3100
 QY 465 AGTTGTCGCAAGTTTATATATCCATTTTATCC-----CAATGTTATCAGTAC 515
 Db 3101 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATCGGACCTGGAAAAATATATGACAA 3160
 QY 516 GAATTAAGATATTTTGACAAAGTTGACAGATTGCAATTTTATGCTCCCAAAAAAGAT 575
 Db 3161 ATATAGAGAGTTATTCATTAATTTAGACAAAGTTGCTCATGTTCAAGCTCAGGTTAAAT 3220
 QY 576 AGGATCGGGTTTGAATTTGCACTGCAATTTATGCTGATTTATATAGAAATTTCA 635
 Db 3221 TGAAGCGGTTTGAATGATGCGGCGGACGATGATCTATCAAGATATGAAATTTCC 3280
 QY 636 GCCAGCTTTGATTAATGACGTTTCAAGTTCTAGAAAGTATCCTGAGAAAGTTCCCAAC 695
 Db 3281 ACCCGCATTAATCTCTAATTTGCGCAATTT-----GGAAGTCTCTTACGGCAATTA 3334
 QY 696 AGAGTTGAAAAAATGATGTAAGTAATCTGGGAATTCAAACATGAAAGATGATCAATACC 755
 Db 3335 ACTGGCGCATTTGGTTGATGTAAGAAAGCTGGAATTTTACGATTAAGTAACATTTACC 3394
 QY 756 ATACGGAATCAAGTTATTAATGAGTGAACGTCAGAGGTTGCTCAGAAACCCCAATTGCT 815
 Db 3395 TTGGGGATTAACTTTATGATGGCGGATTTAAGAAATGTTTCAAGAAACAGTAAATCTGCT 3454
 QY 816 ATACAGGTACTCCAAATGGAAGAAAGAAAGCCAGAAAGAAAGCTCTGTGTATGACCA 875
 Db 3455 CCGAAGGTTAAAAAATTTGTAATGATTCGCATATGCCGAAAGCTTGAAAAATATATACGA 3514
 QY 876 GCTTAATAGTCCCAATTTTACGTTTATGAAGAAATT 911
 Db 3515 ACTGCATCATGCAAAATTTAGATTTATGATGAGACT 3550

RESULT 34
 US-10-835-516-76
 ; Sequence 76, Application US/10835516
 ; Publication No. US20040194162A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hahn, Frederick
 ; APPLICANT: Kuehne, Adelheid
 ; TITLE OF INVENTION: Manipulation of genes of the mevalonate and isoprenoid pathways
 ; FILE REFERENCE: KXS-103XCI
 ; CURRENT APPLICATION NUMBER: US/10/835,516
 ; PRIOR APPLICATION NUMBER: US/09/918,740
 ; PRIOR FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: 60/221,703

; PRIOR FILING DATE: 2000-07-31
 ; NUMBER OF SEQ ID NOS: 76
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 76
 ; LENGTH: 14623
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1..7)
 ; OTHER INFORMATION: Plasmid transformation vector pPhoE, containing Operon E, contai
 ; OTHER INFORMATION: 1
 ; US-10-835-516-76
 Query Match 10.2%; Score 132; DB 18; Length 14623;
 Best Local Similarity 50.5%; Pred. No. 1.5e-22;
 Matches 473; Conservative 0; Mismatches 430; Indels 33; Gaps 5;
 QY 3 GTCAAAAACATTTAGTGCACCTGGAAAAAGCATTTCTGTGTGATATTTGTTCTTGA 62
 Db 2621 GTTGAAGCCTTCAAGTCCCAAGGAAAGGTTACTAGCTGGTGAATATTTAGTTTGA 2680
 QY 63 GCCAATTTATGATCTTATGTCACAGCATTTGTCATCAAGATGCAATGATTAATACAC 122
 Db 2681 TACAAATATGAACATTTGTAGTCGATTAATCGCAAGATGCAATGATGATGATCCATCC 2740
 QY 123 AAAAGAAC-----CAGTTGAAAGATCTAGATCAAAATTTCTTCACCCCAATTGC 176
 Db 2741 TTACGGTTCATGTGCAAGGCTGTGATTAAGTTGAAGTGTGTAAGAAATTAACAATTAA 2800
 QY 177 AAACGGAAATGGGAATATCAATATCATCAATAATACAG--AGAACCCAGAGAATTCA 233
 Db 2801 AGATGGGAGATGGGTACATATTAAGTCTTAAAGTGGCTTCAATTCCTGTTTGCATAG 2860
 QY 234 GTACGGATTAATCCATTTTATAGAGCACTATATTCATGCTTTTACCTTATATCAACC 293
 Db 2861 CGGATCTTAAGAACCTTTCATGTAAGAAAGTTATGCTTAACGTATTTAGCTACTTTAAAC 2920
 QY 294 GACCGAAGCATTTGATCTTGAA-----TCATCATTTTCTCAGACCTCGATATCA 344
 Db 2921 TTAATGAGAGCACTACGCAATAGAAACTTGTGTTATGATATTTCTCTGATGATGC 2980
 QY 345 TTCAACAAGAAATATCTGAACCAAGACATCTCGAATGAGAGAAAAAATTTCTTTTACA 404
 Db 2981 CTACCAATTCAGAGGATAGCGTTACCGAACATCGTGGCAACAGAAATTAAGTTTACC 3040
 QY 405 TTCTCGGCATTAACCAAGTGAAGAACCGGATTAGTTTATCGGCAAGATTAGTGTG 464
 Db 3041 TTGCGACAGAATTGAAGAAAGTTCCCAAAACAGGGCTGGCTCTCGGCAAGTTTACAC 3100
 QY 465 AGTTGTCGCAAGTTTATATATCCATTTTATCC-----CAATGTTATCAGTAC 515
 Db 3101 AGTTTAACTACAGCTTTGGCTCTCTTTTGTATCGGACCTGGAAAAATATATGACAA 3160
 QY 516 GAATTAAGATATTTTGACAAAGTTGACAGATTGCAATTTTATGCTCCCAAAAAAGAT 575
 Db 3161 ATATAGAGAGTTATTCATTAATTTAGACAAAGTTGCTCATGTTCAAGCTCAGGTTAAAT 3220
 QY 576 AGGATCGGGTTTGAATTTGCACTGCAATTTATGCTGATTTATATAGAAATTTCA 635
 Db 3221 TGAAGCGGTTTGAATGATGCGGCGGACGATGATCTATCAAGATATGAAATTTACC 3280
 QY 636 GCCAGCTTTGATTAATGACGTTTCAAGTTCTAGAAAGTATCCTGAGAAAGTTCCCAAC 695
 Db 3281 ACCCGCATTAATCTCTAATTTGCGCAATTT-----GGAAGTCTCTTACGGCAATTA 3334
 QY 696 AGAGTTGAAAAAATGATGTAAGTAATCTGGGAATTCAAACATGAAAGATGATCAATACC 755
 Db 3335 ACTGGCGCATTTGGTTGATGTAAGAAAGCTGGAATTTTACGATTAAGTAACATTTACC 3394
 QY 756 ATACGGAATCAAGTTATTAATGAGTGAACGTCAGAGGTTGCTCAGAAACCCCAATTGCT 815
 Db 3395 TTGGGGATTAACTTTATGATGGCGGATTTAAGAAATGTTTCAAGAAACAGTAAATCTGCT 3454

QY 816 ATCAGAGTACTCCATGAGAAAAAGGAAAGCCAGAAAGAACTCTGTGTTGATGACCA 875
DB 3455 CCAGAAAGTAAAAAATTTGGTATGATTCGCATATGCCAGAAAGCTTGAAAAATATATACAGA 3514
QY 876 GCTTATATAGTCCCAATTTTACAGTTTATGAAAGAAAT 911
DB 3515 ACTCGATCATGCAAAATTTCTAGATTATGATGATGACT 3550

RESULT 35

US-10-369-493-25849
; Sequence 25849, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25849
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-25849

Query Match 6.7%; Score 86.6; DB 15; Length 1269;
Best Local Similarity 56.5%; Pred. No. 1,88-11;
Matches 205; Conservative 0; Mismatches 149; Indels 9; Gaps 2;

QY 499 CCCATGTTATCAGTACGATTAAGATATTTTGGACAACGTTGGACAGATTGGACATTTGT 558
DB 547 CTCAAAATGATGATTAACCAAGTCAATGTTGACATCTGGACAAATTTGCCACTGT 606
QY 559 TATGCCAAAAAAGATAGATCTGGGTTGATCTTGCACACTGCAATTTATGCTGATT 618
DB 607 TCTGCAAGAGAAAGTGAAGTGGTTGATGTTGGCGCTGCCACTTGGGGAAGTTGT 666
QY 619 GTATATAGAAATTTCAAGCCTTTGATTAATGACGTGT-----TCAGTTTCTAGAA 672
DB 667 ATTATATGAAGTTTGAATCTTAATTAATTGAACAAGTGTGCTCCATACGACGAACAA 726
QY 673 AGTATCTGAGAGTTCCCAAGAGTTGAAAAAATTGATGAAGTAATCTGGAAATTC 732
DB 727 ATTAAGAACATTAATTTTCCACGAACTGAGAAAAAATGTTCTAAAAAATG---TCA 783
QY 733 AAACATGAAGATGATACATTACCATACGGAATCAAGTTATTATGAGTGAAGTCAAGGT 792
DB 784 GATGTTGTTCATTTCAATTGCTGCTGACTTATTTGTTGTTATGAGGAGACGTAAGTGT 843
QY 793 GGCTCAGAAACCCCAATTTGATCAAGTACTCAATGAGAAAAAGGAAAGCCAGAA 852
DB 844 GGGAGTAGACACCGGGTATGTAATAAAGTACACCAATGCGAAAGGAAATCCAGAG 903
QY 853 GAA 855
DB 904 GAA 906

RESULT 36

US-10-320-797-2218
; Sequence 2218, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eiroshkin, Alexey M.

APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2218
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Cryptococcus neoformans
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1285)..(1285)
; OTHER INFORMATION: n=g, a, t or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1305)..(1305)
; OTHER INFORMATION: n=g, a, t or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1311)..(1311)
; OTHER INFORMATION: n=g, a, t or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1313)..(1313)
; OTHER INFORMATION: n=g, a, t or c
US-10-320-797-2218

Query Match 5.2%; Score 67.6; DB 16; Length 1506;
Best Local Similarity 48.8%; Pred. No. 1,5e-06;
Matches 209; Conservative 0; Mismatches 213; Indels 6; Gaps 1;

QY 718 AGTAATCGGAATTAACAATGAAGATGATCATTTACCATTCGGAATCAAGTTATTAATG 777
DB 1096 AGCAATGAGATTAACAATCTATCCCTTTCCGACTCCCAAACTTTGCGTCTTACTC 1095
QY 778 GGTGACGTGAAGGTGCTCAGAAAACCCAAATTTGATCAAGTACTCAATGAAA 837
DB 1096 GCCGACGTTTGTGGGTACTGACACGCGCATCTTGTGTCCTCGCTCAAGTGAGA 1155
QY 838 AAGGAAAGCCAGAAAGAGCTCTGTGTATATACAGCTTAATAGTCCAAATTTACAG 897
DB 1156 AATTAAGTATGAGAAAGCCGATGAAGTGGCGAGCTGATTAATGGAACAGGGCC 1215
QY 898 TTATATGAAGATTAAGGGAATGCGTAATAATACGATCAAGCCAGAGACTTATATT 957
DB 1216 TTGGGTAGGTGTGAAGATATGCTTTGAACCTCCCTACACCCGAGCCCTTACGC 1275
QY 958 AAAGATTAAGTATTTCTGTGAGCCTTTGACTGTTGCGATTAAGAAATCAGAAAAAGG 1017
DB 1276 ACATCTC-----CATTTCTTCCACCGGCTGCTTNTGTCATCCGCGCTGTG 1329
QY 1018 TTACAGCATTAACAAAAATCAGAGTTCAATTGAACTGATGCAACCCAGTT 1077
DB 1330 TTGGGTGAAGTGTCCGACTCTGGGGGTACCAATGCAACAAAGAAAGGAAATTTG 1389
QY 1078 TTGACCGTTGTCAAGAGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA 1137
DB 1390 CTGATGATCATGCGGACAAAGTCAAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA 1149
QY 1138 TACGATGC 1145
DB 1450 TACGATGC 1457

RESULT 37

US-10-128-714-2406
; Sequence 2406, Application US/10128714
; Publication No. US20030119013A1

```

; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Broshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128, 714
; PRIOR APPLICATION NUMBER: US 60/285, 697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287, 066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295, 890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303, 899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316, 362
; NUMBER OF SEQ ID NOS: 2001-08-31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2406
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-2406

```

```

Query Match 5.1%; Score 66; DB 15; Length 1356;
Best Local Similarity 45.7%; Pred. No. 3.7e-06;
Matches 547; Conservative 0; Mismatches 595; Indels 54; Gaps 7;

```

```

QY 6 AAAAGCATTAAGTACCTGGAAGAAAGCATTTCTGCTGCTGATATTTGGTTCTTGAGCC 65
DB 24 AACGGCTTTCCTCCGCGCGGCAAGTCTCTCTCCTCCTCCTGATATCTGGTCTCGAGCG 83
QY 66 AATTATGATGCTTATGAGCAGATTTGTCATCGATGATGATGATTAACCAAA 125
DB 84 CAGCTTACCTGGGACGTGTGCTGCTGAGCGCAAGTCCATCGGAGGCAATGTTC 143
QY 126 AGGAACCATTTGAAAGATCTAGATCAAAATTTCTTCAACCCCAATTTGCAAGAGAGA 185
DB 144 TGGAGACAGGAAGAGAGAGAGATGCTGATACCTCCCAAGTTGCTGATGCGGT 203
QY 186 ATGGGAATATCATATCATCAATTAACAGAGAGCCAGAGAGTTCAATCCATTA 245
DB 204 ATGGAGTATATGATCAAAAGATCGAGAGCGAGCTGACCCATCCAGGCCACGC---AA 260
QY 246 TCCATTTTGAAGGCACTATATTCATGCTTTTGTAGTTATAT-----TCAACGAC 297
DB 261 CCGTTTGTGAGAGCTCTCTGAACTTCTGCTTACTTACATCAATGATGCGGACTC 320
QY 298 GAAGCATTTGATCTTGAATATCATATTAATCAAGACCTGATATCAATCAAGAGAT 357
DB 321 GAAGGATTTGGGGTATATGATTAATCTATCTGCGGAGCAAGATTAATCTCCGAGAC 380
QY 358 ACTGAAACCAAGACATCTCGATGAG---AAAAAATTTCTTTTCAATCTCGTGC 413
DB 381 GGCCTTCTCAAGGCTTCGGGACCTCGGTCGTCAAGAGATTCGTGACCTTGTGTTGC 440
QY 414 CATTACGAGTGGAGAAAGACCGGATAGTTTCAATCGGAGATTAATGTCAGTTGTTGC 473
DB 441 CTTTCAAGAGGACACAGAGAGGCTGAGGCTTTCAGCCCTTGTGATGCTGCTGCT 500
QY 474 CACAAGTTTATATCCATTTTA-----TCCCAATGTTATCAATAGATAAGATAT 527
DB 501 GTGCTCTCTGATCAACGCTATATGCAACCGAGATCTCGGTCCAGGCGCGAGCA 560
QY 528 TTGCAACAGCTTGCAAGATGCAATGTTATGCCCCAAAAAAGATAGATTTGGGTT 587
DB 561 GCTTCAATCTGGCCAGGCGGCCACTGCGCTGCCAGGATTAAGTCCGCTGCT 620

```

```

QY 588 TGATGTGCACTGCAATTTATGCTCTGATTTGATATAGAGATT-----TCAGCAGCTT 643
DB 621 CGATGTGCAAGCTGCAATTTAAGGCTCTGCTCTGATCAAGAGCTTCCCTCGATTTCT 680
QY 644 TGATAAATGACGTGTTTCAAGTTTCAAGAAATGATCTTGAGAGTTTCCCAAGATTTGA 703
DB 681 CGATCTGAGGTGAGACCTGCTCTCCAGCTTCAAGAGAGGTTGTTCCGATCGTAGA 740
QY 704 AAAAATGATTTGAAGTA-----ACTGGAATTCAAACATGAAAGATGATACATTACC 755
DB 741 GGAGCGGACCCCGAGATCCGTGGAGACCGAATGCTTGAATTCGATGAGAGCTCC 800
QY 756 ATACGAATCAAGTTATTAATGAGTGAAGTCAAGGCTGCTCAGAAAACACCAATTTGT 815
DB 801 CCGGGAATGAATGATGCTCTGCTGAGAGTGAATGAGTTGCGAGATCTCATCATGAT 860
QY 816 ATCAAGATCTCAATGAGAAAAAGAAAAGCCAGAGAAAGCTCTGTTGTATGACCA 875
DB 861 GAGAAAGTTTGGAGTGGCGGAAACAGAAATCAGAAAGAACCGATATGCTCTGCGGCGC 920
QY 876 GCTTAATAGTCCCAATTTAGCTTATGAGAAATGAGGAAATGGGTGAAATAACA 935
DB 921 TCTGCAATCGAACAAGAGACTTCCCTGGAATCAAGACGCTTGACAGAGCCCGGA 980
QY 936 CTCAGACCCAGAGACTTATATTAAGAGTTAGATCATTTCTGTTGAGCCTTTGACTGTGC 995
DB 981 CGAATATCTCTGATGATCTTGAATATGTCGCACTATATTCAGAGGCTC----- 1032
QY 996 GATTAGAACATCAAGAAAGGTTTACAGATTTAACAACAAATCAAGAGTTTCAATTGA 1055
DB 1033 -----CGTAACCAATCCGTTCCATGATCAAAAGTCGATGCCAATCGA 1079
QY 1056 ACCTGATGTCAAACCCAGTTGTTGACCCCTTGCAGAGATTCCTGCTGTTGTGCTG 1115
DB 1080 GCGCGGCTCAAAACAGACTTGAAGCTCTGCTGAGCTGAGAGGATGTCATCTGCTG 1139
QY 1116 TGTGTTCCAGAGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1171
DB 1140 TGTGTTCCAGAGAGGAGGAGGCTGAGAGGCTGATGATGATGATGATGATGATGATG 1195

```

```

RESULT 38
US-10-032-585-3023/c
; Sequence 3023, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032, 585
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3023
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-3023

```

```

Query Match 5.0%; Score 65; DB 15; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATGCAAGAGATTTAGTGAACCTGGAAGAAAGATTTCTGCTGCTGATATTTGGTTCTT 60
DB 65 ATGCAAGAGATTTAGTGAACCTGGAAGAAAGATTTCTGCTGCTGATATTTGGTTCTT 6
QY 61 GAGCC 65
|||||

```

Db 5 GAGCC 1

RESULT 39
US-10-128-714-7406
; Sequence 7406, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengli
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroskin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7406
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: *Aspergillus fumigatus*
US-10-128-714-7406

Query Match 4.8%; Score 62.6; DB 15; Length 1455;
Best Local Similarity 46.7%; Pred. No. 2.8e-05;
Matches 375; Conservative 0; Mismatches 389; Indels 39; Gaps 4;

QY 387 AAAAACTTTCTTACCATTCCTGTCGCAATTACCGAAGTGGAAGGCGGAGTTAGGTTCC 446
Db 513 AACGAGATTCGTGACTTTGGTGTCCCTTCAGAGGACACAGACAGGCGCTGGGCTC 572
QY 447 ATGGGAGATTGTCAGTTGTCGTCACAGTTTATCCCA-----TTTATCCC 500
Db 573 TTGAGCGCCTTGCTACCTGCCCTGGTGTCTCTGTCATCCACCGTACTATGCAAC 632
QY 501 CAATGTTATCGATCGAATTAAGATATTTTGCACACGTTGCACAGATTGCACATTGTTA 560
Db 633 CGAGCATCTCGTCGACGCGCGGCGGCAAGCTTCAACATCTG6CCGAGCGGCGCCACTGCGC 692
QY 561 TGGCCAAAAAAGATAGGATCTGGGTTTATGTTGCACTGCAATTTATGGTCTGATGTT 620
Db 693 TGGCGAGGCTTAAGTCCGGTCCGCTTCGATGTTGACGTGCGCATTTTACGCGCTCTGTC 752
QY 621 ATATPAGAAGTTTCAGCAGCTTTGATTAATGAAGCTTTCA-----GGTTCTAGAAGTG 676
Db 753 CTACAGAGGCTTCTCCCTCGATTCGAAATCAGTGGTGAAGCGGTGCTTCCAGGCTT 812
QY 677 ATCTGGAAGTTCCCGACAGAGTTGAAAAATTTGATTTGAAGTAACTGGG-----A 728
Db 813 CGAAGAGCGGTTGTTCCGGAATCGTAGAGAGCGCCGACCCAGATCCGTTGGGACACGA 872
QY 729 ATTCAAAATGAAAGATGATGATTCATTAACATACGGAATCAAGTTATTAATGGGTGACGTC 788
Db 873 ATGTCTAGACTTGGCGATGAGCTCCCGCGGGAATCAATGGTCTCTGTGTGACGTGCA 932
QY 789 GGGAGGTCAAGAAACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 848
Db 933 ATGTGTTCGACGATCCATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 992

QY 849 AGAAGAAAGCTGTGTTGTGTATGACCACTTAATAGTCCAAATTACAGTTATGAGGA 908
Db 993 GAAGAAAGCGGATATGCTCTGGGGCGCTCTGCAATGAAACAGAGAGATTCGCGTGA 1052
QY 909 ATTGAGGAAATGCGTGAATTAAGATCTGACGCCAGACCTTATTAATTAAGATTAGA 968
Db 1053 ACTGAGCGCTTGGACAGAGCGCCGAGCAACATCTTCAAGTACTTGAATAATGTCG 1112
QY 969 TCATCTGTGAGCCTTGTGACTGTTGCGATTAAAGACATCAGAAAAGGTTACAGACTT 1028
Db 1113 CACTATATTACAGCGCTG-----CGTAACCATCCGTTCCAT 1151
QY 1029 AACCAAAAATCAAGCTTCAATTGATGATGATGATGATGATGATGATGATGATGATG 1088
Db 1152 GACTCAAAAGTGGAGTGTCCCAATCGAGCGCGCTGCAAAACCGAGCTTATGACGCTCT 1211
QY 1089 TCAAGAGATTCCTGTTGT 1148
Db 1212 GTCCGAGCTGGAGGGGTGATCGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1271
QY 1149 AGCTGATTAGTGTGAAATC 1171
Db 1272 TGTGCTCTCATTCAGACATC 1294

RESULT 40
US-10-128-714-1406
; Sequence 1406, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengli
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroskin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1406
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: *Aspergillus fumigatus*
US-10-128-714-1406

Query Match 4.8%; Score 62.6; DB 15; Length 1509;
Best Local Similarity 46.7%; Pred. No. 2.9e-05;
Matches 375; Conservative 0; Mismatches 389; Indels 39; Gaps 4;

QY 387 AAAAACTTTCTTACCATTCCTGTCGCAATTACCGAAGTGGAAGGCGGAGTTAGGTTCC 446
Db 567 AACGAGATTCGTGACTTTGGTGTCCCTTCAGAGGACACAGACAGGCGCTGGGCTC 626
QY 447 ATGGGAGATTGTCAGTTGTCGTCACAGTTTATCCCA-----TTTATCCC 500
Db 627 TTGAGCGCCTTGCTACCTGCCCTGGTGTCTCTGTCATCCACCGTACTATGCAAC 686
QY 501 CAATGTTATCGATCGAATTAAGATATTTTGCACACGTTGCACAGATTGCACATTGTTA 560

```
Db      687  CGACGATCTCGGTCAGAGCCGCGACAAAGCTTCACAACTCTGGCCAGCGGCCCACTGCGC 746
Qy      561  TGCCCAAAAAAGATAGAGATCTGGGTTTGATGTGCAACTGCAATTTATGGTCTGATGT 620
Db      747  TGCGCAGGGTAAGTCGGGTCGGGCTTCGATGTTCGACCTGCAATTTACGGCTCCTGTCT 806
Qy      621  ATATAGAAAGATTTCAGCCAGCTTTGATTAATGACGTGTTTCA---GGTCTAGAAAAGTG 676
Db      807  CTACAGACGGCTTCTCCCTCGATTTCTGAAATCAAGTGGGTGACGCTGGTCTCCAGGCTT 866
Qy      677  ATCCTGAGAACTTCCCAAGAGTTGAAAAAATGATTGAAAGTAACCTGGG-----A 728
Db      867  CGAAGACCGGTTGTTCCGAGATGTAGAGACGCCGACCCCGACATCCGTGGACACCGA 926
Qy      729  ATTCAACATGAAAGATGATACATTACCAATCGAATCAAGTAAATTAATGAGTGAAGTCAA 788
Db      927  ATGCTTAGACTTCGGCATGAAGCTCCCCCGGAATGCAATGGTCTCTGTGACGTGCA 986
Qy      789  GGGTGGCTCAGAAAACCCCAAATTGGTATCAAGATACTCCAAATGAAAAAGAAAAGCC 848
Db      987  ATGTGTTTCGACATCTCATCGTGAAGAAAGTTTGGAGTGGCGGAAACAGAAATCA 1046
Qy      849  AGAAGAAAGCTCTGTGTGTATGACCACTTAATAGTCCAAATTTACAGTTTATGAAGA 908
Db      1047  GAAGGAAGCCGATATGCTCTGGGGCGCTCTGCAATCGAACAAAGAGACTTCGCTGGA 1106
Qy      909  ATTGAGGGAATGCGTGAAAAATAGACTCAGACCCAGAGACTTATATTAAGAATTAGA 968
Db      1107  ACTGAGACGCTTGGCACHAGGCCGAGACGAACATATCTCACTGACTTTGAAAAATGTCCG 1166
Qy      969  TCATTTCTGTGAGCCTTTGACTGTTCGATTAGAACATCAGAAAAGGGTTACAGCATT 1028
Db      1167  CACCTATATTCAGCGCTCG-----CGTAACCAATCCGTTCCAT 1205
Qy      1029  AACACAAAAATCAGAGGTTCCAAATTGAACCTGATGTCCAAACCAATTGTTGACCGTTG 1088
Db      1206  GACTCAAAAGTCGATGTCCCAATCGAGCCGCGCTCCAAACCGAGCTACTTGAAGCTCT 1265
Qy      1089  TCAAGAGATTCTGTGTGTGTGTGAGTGTTCAGAGTGTGAGTGTGATACGATGCAAT 1148
Db      1266  GTCCGAGCTGAGGGGTGTCACTGGGTGTGTGTTCAGAGACAGGGGGCTACGACGCCAT 1325
Qy      1149  AGCTGATTAAGTGTGAAAAATC 1171
Db      1326  TGTGCTCCTCATCCAAGCAATC 1348
```

Search completed: January 26, 2005, 07:56:03
Job time : 566 secs

This Page Blank (uspio)